Proceedings of the 20th BCS HCI Group conference Volume Two

Fields, Bob; Stockman, Tony; Nickerson, Louise Valgerdur; Healey, Patrick

PROCEEDINGS

of the 20th BCS HCI Group conference in co-operation with ACM

HCI 2006

Queen Mary, University of London

11-15 September 2006

VOLUME TWO

Bob Fields, Tony Stockman,
Louise Valgerður Nickerson, and
Patrick G. T. Healey (Editors)

HCI 2006 Proceedings Volume 2

Bob Fields
Tony Stockman
Louise Valgerður Nickerson
Patrick G. T. Healey
(Editors)

RR-06-07

ISSN 1470-5559

Contents

Preface

Welcome to Volume 2 of the proceedings for HCI2006: Engage!

This volume contains the short papers accepted for and presented at the conference, together with the papers associated with interactive experiences, posters, doctoral consortium presentations, workshops and tutorials. The volume contains a rich mixture of contributions and we invite you to explore them. As usual, we have relied on a group of volunteer reviewers who have assisted the chairs in reviewing the submissions received. We thank them very much for their time and expertise. We had a wide variety of short paper submissions this year, which underwent a rigorous review process. Of the 83 short paper submissions, 36 short papers were accepted for presentation at the conference. For the first time this year, 7 outstanding short papers were selected for publication in volume 1 of the proceedings, the remaining 29 excellent contributions appearing in this volume.

Different themes emerge from these submissions including Enthralling experiences, interactions in the wild, connecting with others, mind body spirit, interactions for me and at the periphery. Topics covered by these papers include mobile interactions, educational software development, computer supported creativity, ubiquitous and wearable systems, designing for social interactions, engaging patients and students, listening to and managing music and the mutual influence of emotion and interaction.

The short papers provide a good representation of the range of issues in current HCI research. Anyone interested in HCI will find a number of thought-provoking ideas to engage with, whether it be a novel interaction technique or an inspiring idea. The papers included in this volume reflect the order of presentation at the conference, but if you're looking for a particular author, there is an author index at the back of the book.

Interaction is clearly at the heart of HCI, and through the Interactive Experiences section of the conference, a number of new and exciting interaction possibilities are demonstrated and explored. To fully appreciate this kind of work, interaction with the technology is required, but the papers published here give a flavour of some new technologies and novel experiences.

The poster submissions similarly represent a wide range of current HCI concerns, including multi-modal interaction, user authentication, novel approaches to usability evaluation, non visual interaction and customisable interface design. Promoting social interaction and supporting diversity are two themes from the short papers that are maintained in the poster submissions.

Each year the conference holds a doctoral consortium, which encourages PhD candidates at various stages of their study to come together and discuss their work. The submissions to this years consortium provide a typically interesting glimpse of a range of new areas of HCI research, combined with new approaches to established problems.

Workshops provide an invaluable opportunity to get to grips with specific problem areas in detail, as well as enabling information exchange between researchers working in the same area. Often the work continues long after the workshop has disbanded. The papers describing the workshops are only an introduction to the area and to the topics being considered. Inclusion is a strong theme of this year’s workshops, with two workshops on designing for older users as well as one on Graduate Career Development for Women in HCI Research.

This year's tutorials combine to provide wide ranging coverage of HCI development issues, from the management of iterative projects, approaches to requirements capture and the use of scenarios, through to the application of specific methods and principles of interaction design

We hope you enjoy this set of papers, and would encourage you to contact the authors if you are interested in any of the topics presented here. Go forth and engage!

Bob Fields (b.fields@mdx.ac.uk)
Tony Stockman (tonys@dcs.qmul.ac.uk)
Louise Valgerður Nickerson (lou@dcs.qmul.ac.uk)
Patrick G. T. Healey (ph@dcs.qmul.ac.uk)
August 2006

Conference Committee

Conference Co-Chairs:
Nick Bryan-Kinns, *Queen Mary, University of London, UK*
Patrick G. T. Healey, *Queen Mary, University of London, UK*

Technical Co-Chairs:
Ann Blandford, *University College London, UK*
Paul Curzon, *Queen Mary, University of London, UK*
Laurence Nigay, *University of Grenoble, France*

Panels:
Olav W. Bertelsen, *University of Aarhus, Denmark*
Adrian Williamson, *Graham Technology plc., UK*

UPA UK liaison :
Caroline Jarrett, *UPA, UK*

Workshops:
Stephanie Wilson, *City University, UK*
Helen Sharp, *Open University, UK*

Tutorials:
William Wong, *Middlesex University, UK*
Paola Amaldi, *Middlesex University, UK*

Short papers:
Tony Stockman, *Queen Mary, University of London, UK*
Bob Fields, *Middlesex University, UK*

Interactive experiences:
Willem-Paul Brinkman, *Brunel University, UK*
Peter Wild, *University of Bath, UK*

Doctoral consortium:
Panos Markopoulos, *Eindhoven University of Technology, NL*
Angela Sasse, *University College London, UK*

Posters:
Dimitrios Rigas, *University of Bradford, UK*

Lab overviews:
Fraser Hamilton, *Designed for All, UK*

Website, submissions, registrations:
Marc Fabri, *Leeds Metropolitan University, UK*
Dave Clarke, *Visualize Design, UK*

Treasurer:
Sue White, *Queen Mary, University of London, UK*

Student volunteers:
George Papatzanis, *Queen Mary, University of London, UK*

Banquet fairy:
Jenn Sheridan, *Lancaster University, UK*

Sponsorship:
Helen Petrie, *University of York, UK*

Publicity:
Paul Cairns, *University College London, UK*

Informal advice:
Ann Light, *Queen Mary, University of London, UK*

BCS Events Chair:
Anxo Roibas, *University of Brighton, UK*

BCS Group Link:
Adrian Williamson, *Graham Technology plc., UK*
Fintan Culwin, *South Bank University, UK*

Previous Conference Chair:
Tom McEwan, *Napier University, UK*

Short Papers

Towards an Ambient Desktop to Promote Workplace Awareness

Rod McCall[1], Benoît Otjacques, Fernand Feltz
Public Research Centre – Gabriel Lippmann
41, Rue du Brill, L-4422 Belvaux, Luxembourg
{mccall,otjacque,feltz}@lippmann.lu

This paper describes a prototype system that displays synchronous workplace awareness information using the ambient technology paradigm. It focuses on making people aware of the nature and volume of interactions that have taken place within a workgroup. The visualisation uses simple graphics based on a maritime flags metaphor, is displayed in the Windows Active Desktop and is designed to be as non-intrusive as possible. The information used in the visualisation is drawn from a database which stores information about the type and frequency of interactions by the users of the co-operative environment. The system uses lightweight technologies and it should be easy to add features to support delivery of the information to standard client applications, smart boards, PDA's and mobile phones.

Workplace awareness, CSCW, visualisation, ambient paradigm

1. INTRODUCTION

Many collaborative tools provide only basic information about the interactions of their users, with information typically being restricted to aspects such as online status (e.g. busy, online or away), how many emails remain unread or when the next meeting is due to take place. While such information is useful it provides only very rudimentary data about the nature and volume of interactions that are taking place. In contrast the work presented in this paper draws upon the experiences of participants within a European Network of Excellence (NoE) and discusses the development of an ambient visualisation (the Ambient Workplace or AW), which informs them about the nature and volume of activities. The objective of the system is to make the participants aware of current activities and to encourage greater participation.

Within projects such as NoE's, information is often shared through traditional means such as email, a web portal, or discussion groups. In many cases the web portal contains associated information such as news, forums or the mailing list archive. In order for people to assess what is new, or the level of activity they often have to break from their existing task and visit the project website, or a collaborative portal. In addition they may find themselves having to use specific tools to take part in discussions or share documents. Examples of such systems include Yahoo Groups or BSCW. Moreover if they wish to assess how active the project, or individual members are they may need to execute a series of search queries on their email or the forums and websites. All of these issues combine to make carrying out tasks which are non-core, e.g. finding out about the level of activity as opposed to writing that all important one hundred page EU deliverable, both difficult and undesirable.

This paper discusses the development of an ambient and ubiquitous visualisation that ensures people are kept aware of the nature and volume of interactions between entities in a co-operative environment. The objective is to provide non-core information in the background without forcing the user to attend to it or use a separate application. The paper contains background information about awareness and the Ambient Workplace, then describes the prototype system, future work and provides a conclusion.

2. BACKGROUND

The model and system developed in this paper rely on two concepts: entities and interactions. Entities are individuals, groups of co-workers (for example a research team), or resources such as documents, email or a shared calendar. Interactions are the actions involved in sharing information among entities, for example sending a document from one person to another.

Awareness is a critical aspect of many CSCW systems and is essential to the overall fluidity and naturalness of collaboration [1]. Ellis argued that 'the philosophy of groupware is to encourage cooperation by making it known and instantly apparent to all who is sharing what with whom' [2]. Awareness of the activities of others also helps to provide a context for one's own activity [3] and may help shape our activities. With this in mind

[1] This work was carried out during the tenure of an ERCIM fellowship.

the AW focuses on providing support for workplace awareness, rather than informal, social or group awareness (for a discussion and definitions of these types of awareness see [4]). Workplace awareness provides information on the identity of those in the workplace, their activities and location.

The AW visualises two kinds of workspace awareness, *Individual Awareness* (IA) and *Global Awareness* (GA). IA provides information about specific entities, for example an individual user or an organisation (i.e. another project partner); examples include the amount of interactions initiated by a specific entity or whether they are online. In contrast, GA provides aggregated or anonymous information about these entities. For example, GA would provide an overview of the number and type of interactions for the entire project, however it does not provide information about a specific person or group of people.

A number of indexes are used in order to calculate and display information relating to the quantity and nature of interactions with respect to GA. A summary is provided below and a more complete description can be found in Otjacques et. al. [5].

- Coopadex (electronic Cooperation Activity Composite Index). This value represents the mean quantity of computer-mediated interactions by members of a group over a given period of time.
- N-Coopidex (Normalized electronic Cooperation Interest Composite Index). This value represents the nature of the co-operation. It provides an indication as to the mean quality or nature of the interactions or more precisely the level of motivation the group members possess for interacting. It takes into account the nature of the task such as whether it is mandatory or optional, active or passive.
- Glocoopex (Global electronic Cooperation Composite Index). This is a global value which is a product of the Coopadex and N-Coopidex values.

The indexes take into account the nature of the interactions which have taken place, for example whether an interaction is active or passive. An active interaction is one you have specifically attended to, for example editing a forum post on a website. A passive interaction being one a person does not need to deliberately attend to, for example noticing that other people are logged into the website. Other categories used are whether the interaction is mandatory or optional, examples include whether a person must reply to an email from their manager or whether they can ignore the email about the office football team.

Although the system makes use of historical data in order to calculate the indexes, for example to calculate mean scores for a period of time, it provides up to date (or synchronous) information. For example a change to the level of activities will result in the visualisation being updated. In order to support GA, the system visualizes the current value of the indexes. Historical values may be valuable in some contexts but these go beyond the scope of this paper.

To support IA, the system draws on data from the entities which the user is monitoring. For instance, whether a user is online, whether a user has initiated interactions in a recent past or whether a given document has been updated.

3. INITIAL PROTOTYPE

3.1 Design Rationale

The Ambient Workplace builds on the idea of peripheral awareness introduced by Heath and Luff [6]. It draws on a number of design principals outlined by Cadiz et. al. [7] which were used in the Sideshow system, namely:

- *Always present: the visualisations exist within standard applications such as email clients, word processors, instant messaging clients and the Windows Active Desktop. This should reduce the need for people to use specific applications in order to obtain awareness information.*
- *Minimize motion: none of the visualisations make use of data or animations with a high frequency of update.*
- *Make it personal: users will be able to customise what they view and other aspects, such as update time etc.*
- *Extensible: our definition is slightly different from that used by Cadiz et. al however plug-ins will be available for a range of standard client applications.*

Others such as making it scalable or supporting quick drill down are not quite as relevant in the context of this system. Cadiz et. al talk of scalability in the context of new sources of information becoming available (e.g. a news feed). With the AW the main changes to data are likely to be the availability of new flags or the addition of new users. Both of these are comparatively easy to handle. Drilldown is also not such a relevant issue as data is unlikely to be very hierarchical in nature, indeed the most complex drilldown required may be when a person wishes to open a document related to a specific entity, again this can be handled relatively easily.

3.2 Implementation

A prototype was developed (see figure 1) to explore the graphical metaphor (see section 3.3) and to test the underlying technologies. The system consists of three main components a server application which retrieves information from a database, a client on the user's PC which displays the visualisation in the Windows Active Desktop and a Jabber Instant Messaging server. The latter is used to control the user accounts (such as log in, presence and privacy information) as well as to send information to and from the client and server applications. The database stores the information about the nature and level of interactions, this is obtained from monitoring the mailing list, website and other sources.

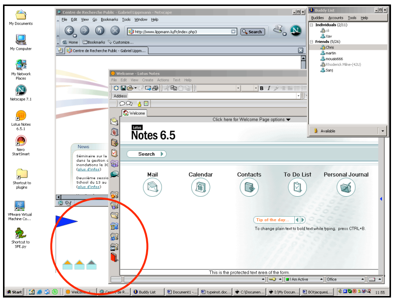


FIGURE 1: A screenshot illustrating the Ambient Workplace (lower left in circle)

3.3 Metaphor

The system uses representations that are based on maritime flags and beacons found on ships and around coastlines. The designs chosen draw upon the basics of the maritime flag metaphors, i.e. the use of different colours and shapes but do not use the same graphics. The metaphor was chosen as many people are familiar with the basic concepts of flags and beacons, and the representations used are comparatively simple. In keeping with the ambient paradigm the aim of the flags is not to distract the users, hence they change comparatively slowly and the amount of data contained in each representation is quite small. Hence once understood they should not require the user to spend a lot of time trying to interpret them. Indeed the intention is that they sit in the background and require almost no specific attention.

Global Awareness is represented by a blue flag (see figure 2). The colour of the flag provides an overview of the quantity and quality of the interactions (Glocoopex), with brighter blue indicating a higher value, and dark blue a lower value. The width of the flag represents a value relating to the use of the cooperation tools (Coopadex), and the height of the flag indicates the level of interest all users had with respect to the interactions (N-Coopadex). Note height relates to the vertical size of the flag, not its position on the flagpole.

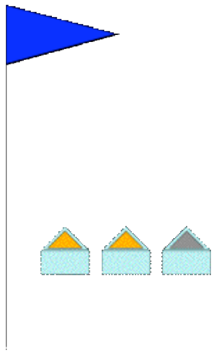


FIGURE 2: A picture of the Ambient Workplace. From the top left, the *Glocopex* flag, the three beacons illustrating entities using the system

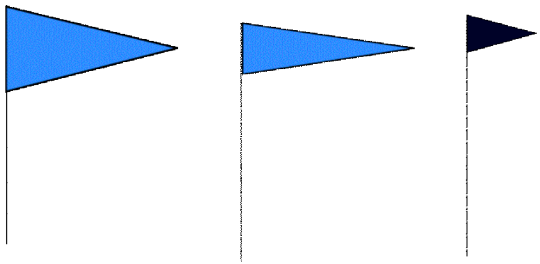


FIGURE 3: Three activity flags indicating the volume and type of interactions taking place within the system

The examples in figure 3 illustrate a number of scenarios based on the values of the *Glocoopex*, *Coopadex* and *N-Coopidex* values. The first flag (figure 3, left) illustrates that there has been a large amount of activity relative to the preset time period (the flag is bright blue), there has been a large amount of voluntary activity (the flag height is high) and there has been a large use of cooperation tools (as represented by the length of the flag). Figure 3 (centre) indicates a slight variation, with there being less interest in the interactions, as represented by the flag height and figure 3 (right) indicates there has been little in the way of voluntary activity and use of electronic cooperation tools. This is represented by the flag height being less, the colour being dark blue and the length of the flag being lower.

FIGURE 4: The Entity Beacons, the further to the right (red) the more interactions or activities that entity has initiated

Individual Awareness is represented by a set of beacons (see figure 4), with each beacon representing an individual or group of individuals (e.g. a project partner location). The coloured triangle indicates the number of computer-mediated interactions concerning that entity over a preset time period. Note that interactions include both active operations like sending an e-mail and passive operations like being updated for a shared resource. Grey indicates no interactions, orange some interactions and red many interactions. At this stage the values are preset, however they can be changed or in future set by the user. When a user or group is online the border of the triangular area will become green (this is not currently implemented or illustrated). This feature only applies to entities that can initiate some interactions (i.e. 'users').

4. CURRENT WORK AND FUTURE DIRECTIONS

An extensive study is currently ongoing with the objective of improving the design of the visualisation. During the first phase around twenty participants (some being potential users) took part and they were asked to draw a series of visualisations using the flags and beacons. A second phase will involve a focus group who will create a shortlist of their preferred designs and if necessary improve them. The third phase will involve as many members of the NoE as possible selecting from the short list which designs they would like to see implemented in the system. The remaining testing will become an iterative process where a series of live prototypes will be developed, evaluated and improved.

As the system develops new features will be added, such as the ability to click on parts of the visualisation to display the relevant data or application. Other features such as visualisations representing the status of email and the news forums will also be added. The intention is also to implement all or part of the visualisation in other client applications, on mobile phones and on smart boards.

5. CONCLUSION

This paper has discussed the development of a system to promote workplace awareness. The system uses the ambient technology paradigm to deliver non-core information to users, with the objective that they can obtain the information without interrupting their normal working patterns. An early working prototype has been developed to explore the underlying ideas such as using the Windows Active Desktop to display information and the maritime flags metaphor. Further work is planned to refine the design of the flags and to add new features.

REFERENCES

[1] Gutwin, C. and Greenberg, S. (1999) The effects of Workspace Awareness Support on the Usability of Real-Time Distributed Groupware. ACM Transactions on Computer-Human Interaction, Vol. 6, N° 3, pp.~243-281.
[2] Ellis, C. (1995) Keepers, Synchronizers, Communicators and Agents. ACM SIGOIS Bulletin, Vol. 15, N° 3, pp.~ 10-14.
[3] Dourish, P. and Belloti, V. (1992) Awareness and coordination in shared workspaces. Proceedings of CSCW'92. Toronto, Canada, October 31 – 4 November, pp.~107-114. ACM, New York.
[4] Greenberg, S., Gutwin, C., and Cockburn, A. (1996) Awareness Through Fisheye Views in Relaxed-WYSIWIS Groupware. Proceedings of Graphics Interface Conference 1996, Toronto, Canada, May 22-24.
[5] Otjacques, B., Noirhomme M., Figueiredo, J., and Feltz, F. (2006), Composite Indexes as Metrics of Cooperative Activity. Proceedings of AIM 2006, Luxembourg, June 7-9.
[6] Heath, C. and Luff, P. (1991) Collaborative activity and technological design: task coordination in London Underground control rooms. Proceedings of ECSCW'91, Amsterdam, The Netherlands, September 24-27, pp.~65-80.
[7] Cadiz, JJ., Venolia, G., Janke, G. and Gupta, A. (2002) Designing and Deploying and Information Awareness Interface. Proceedings of CSCW'02. New Orleans, LO, 16-20 November, pp.~314–323. ACM, New York.

Blowtter: a Voice-Controlled Plotter

Sama'a Al Hashimi
Lansdown Centre for Electronic Arts, Middlesex University
Hertfordshire, United Kingdom
Samaa.alhashimi@gmail.com

The use of microphones as an input device has predominantly represented a linguistic rather than a paralinguistic gateway to computer-based technologies. However, speech is not the only channel of communication that the voice may employ. Moreover, the output of many voice-controlled applications tends to be confined to on-screen displays. By contrast, this paper investigates the combination of non-verbal vocal input with physical outputs. The paper reviews existing work which exploits non-verbal vocalisation, and discusses the relative merits of the verbal and non-verbal as inputs. It presents the *mic-board,* an attempt to find a new application for the microphone. The *mic-board* is used to control *Blowtter*; a voice-controlled plotter. The main aim is to introduce the concept of a *voice-physical* installation and to explore the use of the non-verbal channel of voice as a complementary input mechanism to the verbal channel.

Voice, paralanguage, plotter, voice-physical, speech-recognition, non-speech recognition, mic-board

1. INTRODUCTION

The availability of a wide variety of input mechanisms today reflects the high potential of the human body as a rich source of input. Recently, voices, fingers, hands, and eyes have all been explored and employed in different multimedia applications as an input. This has, to a certain extent, transformed part of the computer monitor into a mirror that reflects the user sitting in front of it while also displaying the visuals processed behind it. A lot of these techniques, however, are not enough to break the barrier between the user and the computer. The widespread reliance on the visuals, and hence the desktop screen holds many users back from the perception of an amorphous computer that any device may incorporate or embody. Only recently has more attention been driven towards what O'Sullivan and Igoe refer to as "Physical Computing" or using computers to affect the physical world:

> When asked to draw a computer, most people will draw the same elements: screen, keyboard, and mouse. When we think "computer," this is the image that comes to mind. In order to fully explore the possibilities of computing, you have to get away from that stereotype of computers. You have to think about computing rather than computers. Computers should take whatever physical form suits our needs for computing. [8]

With the increasing ubiquity of computer vision and video tracking, physical movement has almost always been associated with and expected from the user. When expected from the computer, physical movement is often linked with android (humanoid robot) behaviour either in reaction to a remote control or to spoken commands. In this paper, however, I aim to reflect my interest in programming computers to *complement* users' deficiencies rather than merely to imitate their abilities. *Voice-visual* applications, for instance, may counterbalance our inability to measure voice characteristics and produce visuals in reaction to them. Igarashi and Hughes refer to the use of non-verbal voice to control interactive works as "voice as sound" [4]. On the other hand, our inability to naturally use voice to physically control and affect real inanimate objects calls attention to the development of *voice-physical* applications. It is important that developers adequately exploit the computer's capacity to respond multimodally to voice not only by visuals but also by movements, odours, or any kind of physical output. According to Levin, a computer may cause "uniquely ephemeral dynamic media to blossom from the expressive 'voice' of a human user" [5]. *Voice-physical* installations are likely to prove a fruitful expansion of the possibilities already shown to be inherent in voice-visual and other existing forms of voice-controlled installations. Such an expansion seems to have a natural synergy with the move towards "The Invisible Computer" described by Norman [7].

In light of this, I was prompted to investigate the possibility of programming voice to control inanimate objects in what could be referred to as *vocal telekinesis*. An initial development in this direction was *sssSnake*; a *voice-physical* version of the classic 'Snake' game. In this two-player game, one player makes an 'ssss' sound to control a virtual snake projected on the surface of an installation table. The other player makes an 'ahhhh' sound to move a real coin placed on top of the table. The 'sss' and 'ahh' sounds are distinguished through frequency-range differentiation rather than speech recognition. The game is based on utilizing the difference in frequency between the high-pitched 'ssss' and the relatively low-pitched 'ahhh'. The position of the players round the table, which is fitted with a microphone on each of its four sides, determines the direction of the coin or snake. A plotter is hidden below the table. Its head, to which a magnet is attached, is programmed to move away from the source of the 'ahhh' sound and pull the coin with it. The development and usability testing of *sssSnake* has inspired and paved the way for the development of *Blowtter* which will be discussed thoroughly in the rest of this paper.

2. BLOWTTER

Ali Abbas, an Iraqi orphan whose arms were amputated following an attack in Baghdad, was later fitted with prosthetic limbs. Later on, he started drawing palm trees and temples with his feet [2]. I remembered him when the plotter that I was testing for *sssSnake* started moving in reaction to voice. This led to the development of *Blowtter*.

2.1 Concept

Blowtter is a voice-controlled plotter that allows a disabled user to blow into a *mic-board* in order to draw. The *mic-board* is a small square board that consists of four microphones, one on each of its four sides. Blowing into one of the microphones moves the head of the plotter in the opposite direction as if pushing it. The notion of *blowing* fits neatly with the action of *pushing* the head. Blowing is used as an input in order to facilitate the directness and continuity required for drawing which is not easily achieved by using a spoken command repetitively such as saying "move…move…move" or "move, 10". Igarashi and Hughes further explain the advantage of this technique:

> […] one can say "Volume up, ahhhhhh", and the volume of a TV set continues to increases while the "ahhh" continues. The advantage of this technique compared with traditional approach of saying "Volume up twenty" or something is that the user can continuously observe the immediate feedback during the interaction. One can also use voiceless, breathed sound. [4]

Where appropriate, speech commands are also used in *Blowtter*. For example, "start plotter" and "end plotter". Saying "up" raises the head and allows for moving it without drawing. Saying "down" moves the pen into the paper. Saying the number of the pen causes it to be selected.

The advantage of using blowing lies in the possibility of targeting it with one microphone without interference through nearby microphones. Blowing is perceptually voiceless despite the fact that the microphones may detect it and the computer may measure its voice characteristics. Unlike voiced sounds which involve the vibration of the vocal folds while being generated, blowing and other unvoiced sounds only involve air passing through the larynx without causing the vocal folds to vibrate. Furthermore, the range of amplitude values that voiced sounds can produce is wider and louder than the range of amplitude values generated by blowing. This fact is exploited in specifying a higher threshold for speech than for blowing. This will be explained in detail in the following section.

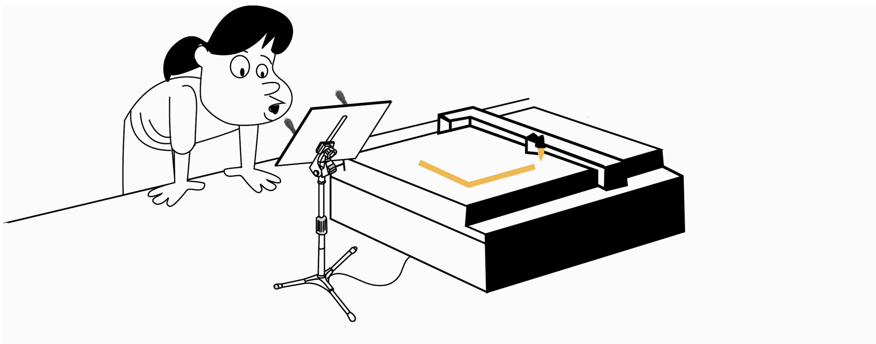


FIGURE 1: *Blowtter*: a voice-controlled plotter that allows a perhaps disabled user to blow into a *mic-board* in order to draw

2.2 Technical Details

The main hardware requirements for the development of *Blowtter* include a DXY-880 Roland plotter, four USB audio adapters, four microphones, and a fast computer. The application has been programmed in Macromedia Director/Lingo. Three Xtras (external software modules) are used: asFFT, Chant, and Direct Communication.

2.2.1 Using asFFT Xtra to detect voice characteristics

The vocal signal is analysed using the Fast Fourier Transform (FFT) algorithm which the asFFT Xtra employs. The Xtra transforms the voice signal into frequency-based data. Through programming it to measure and compare the different volume levels at each microphone, the software recognizes which microphone of the four the user is blowing into. The microphone at which the maximum amplitude is detected is the only one that the software uses to detect voice and to determine the direction for moving the pen. In other words, even when the user blows hard enough for nearby microphones to pick up the sound, the software is programmed to compare the amplitudes and to only react to the microphone through which the loudest amplitude is detected. To avoid the head reacting to any voice input other than blowing, an amplitude threshold is specified in order not to allow noises above it to move the head. When a volume above the threshold is detected, the software assumes that the user is speaking rather than blowing and the speech recognition Xtra, Chant, is activated. Furthermore, to avoid the head reacting to soft ambient noises, another lower threshold is specified. Any noises below it are completely ignored. The Hewlett-Packard Graphics Language (HPGL) was used to program the plotter to move its head in response to blowing.

2.2.2 Using Chant Xtra to recognize speech
Chant Speech Kit, a speech-recognition Xtra for Director, is used to recognise the commands spoken by the user. The Xtra supports a number of speech recognition applications, in this case Microsoft SAPI. It consists of a command, grammar, and dictation vocabulary list. For *Blowtter*, however, only the command list is enabled and customized to consist only of the commands that control the plotter. Thus, detection errors are minimized.

2.2.3 Using Direct Communication Xtra to establish a parallel or serial connection with the plotter
Direct Communication Xtra allows Director to communicate with an external device either through the serial or, as in this case, the parallel port. Care must be taken to match the settings at the computer and at the plotter end.

3. THE MICBOARD AS AN ALTERNATIVE INPUT MECHANISM

The investigation of paralinguistic vocal control of interactive media does not only involve singing, humming, or making voiced sounds as ways to control a computer. Blowing has also been utilized innovatively in very few interactive works. The *Yacht* game by Nintendo DS, for instance, allows the player to blow into the microphone to propel a boat and steer it in the sea. The weight of breath determines the speed of the yacht. Another Nintendo game, *Candles*, also allows the player to blow into the microphone to blow out candles; the larger the size of the candle, the harder the player must blow. In both cases, users probably believe that the velocity of their breath is the determining factor, whereas it is really the volume of the sound created by blowing on the microphone.

Blowing "Windows" by Matsumura at The Royal College of Art, London, is another application that employs blowing [6]. It allows the user to blow into one end of a hose of which the other end is directed towards a computer's screen in order to control, move, and rearrange desktop icons. The duct contains a wireless microphone that detects the blow and measures its intensity. This intensity as well as the size of the file represented by the icon determines the speed at which the icon moves. The hose also contains tilt switches that detect the angle at which the hose is held and determine which side of the desktop to rearrange.

Another remarkable application of blowing is *Kirifuki* which allows for interaction with visual desktop objects by inhalation and exhalation. The system consists of a breath microphone switch containing a Polhemus sensor that detects the orientation of the user's head, a projector that projects the desktop on a desk, and a magnetic gyroscope [3]. The microphone differentiates between inhaling and exhaling by means of comparing between their acoustic signals. Iga and Higuchi claim that the randomness of the exhale signals is less than that of the inhale signal, and by measuring and comparing the local peak numbers the system recognizes the sound input as either an inhale or an exhale [3]. The implementation of this technique allows for a variety of interaction mechanisms. The technique can be applied to make a blow diffuse the icons around the mouse pointer, while an inhale may assemble these icons again. Another interesting application is programming the blow to cut a virtual object, while an inhale pastes it again. If the object being manipulated is three-dimensional, an inhale deflates it while an exhale inflates it. The system may also be used for drawing where an exhale sprays a shape while an inhale erases it.

One last noteworthy installation is *Blow Up*. Developed by Scott Snibbe, this installation allows the sender to blow into a group of twelve small impellers that control an array of twelve large fans which replicate and magnify the speed and movement of the small impellers on the receiver's end [9]. The array of small impellers is placed in one side of the gallery, while the fans they control are placed in another side. Through the development of this installation, Snibbe seems to emphasize that the visual and physical impact of breath on nearby objects can be more obvious than the audible impact. He believes that the physical interaction of our voices and breaths with surrounding media has a significant influence on our inference about the existence of these body activities [9].

In addition to the exploitation of blowing as an input source, *Blowtter* involves a utilization of the un-voiced-ness of blowing to propose a new use for the microphone. The fact that blowing is un-voiced makes it possible to place the four microphones used to control *Blowtter* in a very close position to each other on the *mic-board*. The size of the board is around 12x12 cm and the distance between one microphone and the other is around 11 cm. This minimizes the need to move from a microphone to another where only the user's face is expected to move slightly to direct the air stream into one of the microphones. For such an application designed especially for disabled users, minimizing movement is very necessary. The *mic-board* may be used by disabled users as an alternative to a joystick in some modified versions of existing games. Further implementations may involve increasing the numbers of microphones. Hence, several microphones may be placed on a circular rather than a square board or even forming a matrix of microphones. Blowing into a microphone on the right side of the board and then into a microphone on the left side would draw a straight line from right to left. Moving the head slightly in a certain pattern while blowing sequentially into one microphone after another would draw the pattern or shape that the user's head composes while moving. This would increase the opportunity of allowing smoothly graduated control.

4. CONCLUSION

The development of *Blowtter* is still in progress and user-testing it will certainly reveal many unexpected results. One of the main aspects that I aim to investigate while testing it is the integration of certain paralinguistic components as a complementary input mode to certain speech input applications in order to create a synergistic combination that might let the strengths of each mode overcome the weaknesses of the other. One possible outcome could be a *voice-physical* avatar or robot that is pre-programmed to recognize paralinguistic and linguistic utterances and physically display their equivalent in sign language. This robot may be portable or may also be connected to or placed next to a television for the hard of hearing. Comparing this interactive work with other voice-controlled applications that only involve non-speech input will make it possible to research further paralinguistic voice-related dimensions. One facet that requires further research, for instance, is the use of non-speech voice as an alternative input in comparison to the use of it as a complementary input mechanism to speech recognition.

On the basis of an empirical evaluation of previous work [1], it seems more appropriate to employ paralinguistic input alone to control a game or an entertaining work than to control a practical and serviceable application. Interaction with a game allows for less accuracy than for a functional application. A game usually offers a chance to try again after losing. Error and loss are meaningful components of the action in most games but surely not part of any practical application. For this reason, a game is especially likely to benefit from exclusive use of paralinguistic input, whereas in the case of a practical everyday application, this kind of input is likely to be of greatest benefit when used with a complementary input mechanism.

One advantage of using blowing in *Blowtter* rather than using voiced sounds as in *sssSnake* is that it ensures accuracy by minimizing microphone interference. The underlying technique is shared in both applications, but the extent to which a technical inaccuracy is perceptible and tolerable by the user is not. Because of a limitation of asFFT when used with multiple microphones, two microphones could wrongly detect the same amplitude level even when the player was much closer to one microphone than the other while making loud sounds in *sssSnake*. Occasionally, this complicated the amplitude comparison operation and caused the coin to momentarily deviate. That was, nonetheless, hardly noticeable by the players who were, anyway, running and laughing. For a disabled user sitting still and using *Blowtter* to draw accurately, however, such deviation is intolerable and unquestionably perceptible. Therefore, the un-voiced-ness of blowing and the inherent notion of having to bring the mouth very close to the 'blown' object, which in this case is one of the microphones, make it possible to obviate this problem.

Another important aspect to keep in mind while developing a voice-controlled work is its context. Since fun and laughter are part of an entertaining system, making non-speech sounds to control a game is reasonable and even attention-grabbing to passers-by. On the other hand, a user interacting with a practical voice-controlled interface will most probably not want an audience watching and laughing. As *Blowtter* is not designed as a game, this is another reason behind choosing blowing rather than vocalizing to control it.

The use of *Blowtter* as a drawing tool for physically impaired users could have significant potential. Non-speech voices can provide this user segment with a new dimension for exploring their artistic talents and communicating their thoughts. Exploring the untapped dimensions of voice-control might empower them in other ways too.

REFERENCES

[1] Al Hashimi, S. (2005) Beyond Using Voice as Voice. In Proceedings of the 16th International Conference for Advanced Studies in Systems Research, Informatics and Cybernetics. Baden-Baden, Germany

[2] Gordon, Y. (2006) Ali has the World of Art at his Feet. Kingston Guardian (http://www.kingstonguardian.co.uk/display.var.704393.0.0.php).

[3] Iga, S. and Higuchi, F. (2002) Kirifuki: Inhaling and Exhaling Interaction for Entertainment Systems. Transactions of the Virtual Reality Society of Japan TVRSJ.7 No.4, 445-452.

[4] Igarashi, T. and Hughes, F. (2001) Voice as Sound: Using Non-verbal Voice Input for Interactive Control. Proceedings of the 14th Annual Symposium on User Interface Software and Technology, ACM UIST'01, 11-14 November, pp.155-156. Orlando, Florida.

[5] Levin, G. (1999) Interface Metaphors and Signal Representation for Audiovisual Performance Systems (http://acg.media.mit.edu/people/golan/thesis/proposal/).

[6] Matsumura, E. (2005) Blowing 'Windows'. (http://www.interaction.rca.ac.uk/alumni/04-06/Eriko/html/01_puppet.htm).

[7] Norman, D. (1999) The Invisible Computer: Why good products can fail, the personal computer is so complex and information appliances are the solution. MIT Press, Cambridge.

[8] O'Sullivan, D. and Igoe, T. (2004) Physical Computing: Sensing and Controlling the Physical World with Computers. Thomson Course Technology PTR, Boston.

[9] Snibbe, S. (2005) Blow up. (http://www.snibbe.com/scott/breath/blowup/index.html).

Ubiquitous Inspiration: a Field Study of Artists and Creative Environments

R. T. Jim Eales[1], Dharani Perera[2], Sophie Nichol[2]
[1]School of Computing Science, Middlesex University, London NW4 4BT, UK.
[2]School of School of Engineering and IT, Deakin University, Geelong, Victoria, Australia.
j.eales@mdx.ac.uk

There is a need for a greater understanding of how information technology can best support the creative user. In this paper, we consider the possible influence of the environment on creativity and how this might in the future be supported by ubiquitous computing technology. We present a number of case studies of artists, focusing in particular on their different creative environments and what inspires them. Our findings suggest that artists can be grouped into those that are internally inspired and those that are inspired by external influences. This division suggests a connection to the influence of micro and macro environments. We attempt to derive a number of insights on creativity and creative environments from our investigations.

Creativity, creative environments, computer-enhanced creative environments, ubiquitous computing

1. INTRODUCTION

We consider the technological support of creativity to be a particularly important and interesting area for research. For more and more people, creativity is playing an increasing part in their working lives as well as in their leisure activities. In HCI, designing for the creative user is likely to become an increasingly important issue [2]. Creative industries are expanding rapidly. Indeed, Florida [6] argues that creativity is the driving force of economic growth. He suggests that regions and countries that do not embrace creative communities – in all their various facets and dimensions – are doomed to economic stagnation.

Researching the area of computer-supported creativity is largely a case of exploring an uncharted territory. There are few theories, studies or even concepts to guide us. The study, outlined in this particular paper, was inspired by a quote from Czikszentmihalyi, "It is easier to enhance creativity by changing conditions in the environment than by trying to make people think more creatively." ([3], Pg. 1). Candy and Edmonds [1] have also suggested that outstandingly creative people seem to be able to arrange for their own creative conditions to be available. The development or enhancement of creative environments may perhaps be the best way to approach the difficult task of technologically supporting or stimulating creativity. And perhaps the most productive way to apply or utilise computing power and resources is to embed computers in our natural movements and interactions with our environment, so that we can seamlessly interact with a technologically-enhanced "world at large", i.e., the ubiquitous and pervasive computing paradigm [9]. If we merge these two perspectives, we can conceive of *computer-enhanced creative environments*; Spaces that support creative thought and activity by embedding computing power within the general fabric of the environment and the artefacts associated with an environment. In this paper, we set out to explore the general prospects for computer-enhanced creative environments.

2. CREATIVITY SUPPORT SYSTEMS

Our long-term research plan focuses on the information technology (IT)-based *support* of creativity [8]. We emphasise the term support because we are certainly not attempting to develop software or hardware *incorporating* creativity. We do however believe that IT has an important potential role as a tool or system to support people working in many different areas of creative endeavour. Our specific objective is the design and development of what we term *creativity support systems* (see figure 1). This diagram, our model, although simple, has proved a valuable aid in this uncharted research territory. There are many avenues to explore, in this paper, we consider the possible application of technology to creative environments, and because we are considering artists, our focus is mostly on individual environments (the top left quadrant of our diagram). However, we are also interested in a whole range of technological applications, from environments to special-purpose tools. We are also interested in a wide range of creative practice from very individual instances of creativity to examples of collective creativity.

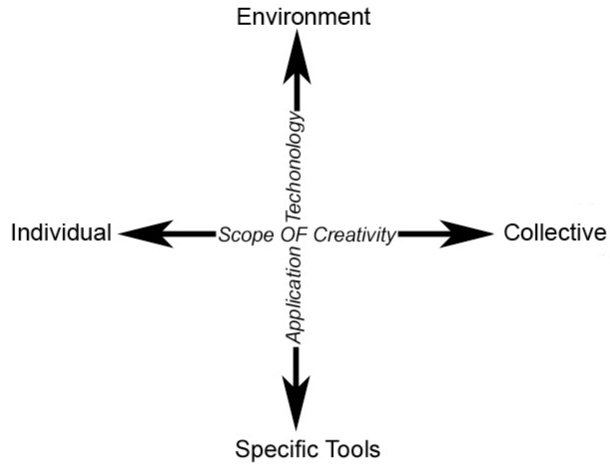


FIGURE 1: Creativity support systems

To advance our general understanding of this research area we have initially set out to study examples of creativity in action in various natural settings. An important part of our investigation is to ascertain the particular needs of the creative person. Although people's creativity may be very different, our hope is that the attributes of IT that creative users find useful may be reasonably constant and generalisable. At this stage in our research, we are primarily studying the creative practice of acknowledged artists. Even at the best of times creativity is hard to find. We consider artists to be specialists in creativity, normally with clearly defined creative output and well-developed and largely stable creative processes, so artists offer a number of advantages when studying creativity in action.

3. STUDYING CREATIVITY

In spite of considerable research effort, human creativity is still not fully understood. Mayer ([7], p. 458) suggests that "although creativity researchers have managed to ask some deep questions. They have generally not succeeded in answering them." However, there does seem to be a general consensus that the two defining characteristics of creativity are originality and usefulness or value. But there is no single indisputable theory or model of creativity that can form the basis for the design of technology to support creativity. In this study, we focus on two related research questions:

- How are creative people inspired?
- What makes a creative environment?

Our ultimate and still remote research question is how might computing power best be used to enhance creative environments? In an attempt to answer these questions we conducted a number of contextual interviews with a variety of artists. The artists described are real people and the names used are their real names. The interviews were tape-recorded and later transcribed. As we were particularly interested in creative environments, we also engaged in observation whenever possible. These observations were recorded as notes, diagrams and digital photographs. We chose to focus on traditional rather than digital artists partly because of opportunity and partly because we believe that the influence of environments and tools is easier to observe in the creation of non-digital art. We also plan to study digital artists.

4. CASE STUDY 1 – JILL

Jill is a full-time Australian artist working in the traditional medium of paint on canvas. She regularly exhibits in Melbourne, Sydney and Brisbane. Jill's paintings are generally large colourful canvases depicting primitive animal characters and human figures. These images represent people and situations from Jill's memories, dreams and relationships. This undoubtedly has implications for her general creative practice since she needs to create the conditions for these images to emerge from her conscious and unconscious mind. Jill has created a large studio in her house. Jill uses her computer and peripherals, situated in a corner of the studio, to assist her in her painting. We have described Jill's creative process in more detail in previous publications [4],[5]. In our study of Jill's interesting use of digital technology, two significant and separate areas were evident. Sometimes, when she is in need of fresh inspiration, Jill will investigate ideas for new paintings using her computer. We have described this as *electronic collaging*. She may spend a week on the computer just playing with ideas. She generally starts by collecting a series of mostly random scanned images. For example, she will open a magazine at a random page and scan that page. She will also scan random images from books, her paintings, her sketchbook and even physical objects such as leaves. These images, or more likely just parts of these images, will be arranged into collages on the computer screen. If she particularly likes an image generated on the computer this will form the basis of a new painting.

Sometimes during the painting of a canvas, switching from paint on canvas to a digital representation offers her a number of advantages. We have termed this *media switching*. To achieve this she takes a photograph of her painting using a standard digital camera. She uploads this image into image editing software on her computer. Jill then uses the software tools to work on the digital image in various ways exploring possible compositional changes.

In our previous studies of Jill's use of digital technology we became aware of a significant split between her initial *inspirational* and subsequent *functional* needs. Although such divisions are never clean cut, when Jill uses her digital technology for electronic collaging she is essentially fulfilling an inspirational need, whereas when she does media switching she is primarily making use of the functional digital advantage of experimentation with conservation. This case study particularly focused our attention on the importance of inspirational rather than functional factors in the development of creative environments. In addition, when Jill is using the computer for inspirational purposes she attaches great value in the ability of the computer to introduce (usually accidental) *randomness*, chance and choice as she refers to it. One of the most significant ways that computing power may be used in the support of creative inspiration may well be to introduce randomness or chance into an environment.

5. CASE STUDY 2 – FAYE

Faye is a trained visual artist living in Melbourne, Australia. She creates mainly landscape and cityscape paintings using oil or acrylic paint on canvas. Although Faye has a small purpose-built studio at her home, most of her paintings are produced in an outdoor setting. For the purposes of this research we observed Faye both in the field and in her studio. Faye particularly looks for artistic inspiration in the physical environment. She is always looking for a scene to paint. "I really like the outdoors. The light, clouds, rain, the atmosphere, people, it's got to look alive; that's what I look for." In the country, Faye looks for interestingly shaped trees, different colours of the leaves, the shape of the terrain, different weather conditions, and so on. In the city she likes to paint old buildings, lampposts and trams. In the country, Faye usually sets up her easel away from her car. In the city, she normally uses her car as a mobile studio. She uses her car to carry her tools such as paint, easel and brushes, but also uses the tail-gate as cover when painting in adverse weather conditions and for privacy. When painting in the outdoors, the only significant distraction Faye faces is strangers that ask her questions about her painting. Faye has a small purpose-built studio at the back of her house. The studio has little natural lighting; so artificial lighting is usually required. The walls of the studio are decorated with her own and other paintings. She usually listens to the radio and television when working in the studio. This studio is mostly used for finishing off landscapes that were started in the field. She often uses photographs to add detail to these paintings. Things that distract Faye when painting in her studio are the telephone and friends or family members dropping-in for visits.

Our study of Faye showed even more clearly the split between *inspirational* and *functional* artistic needs. Since the physical environment is her primary source of inspiration, she has to transport her functional requirements to the site of her inspiration. Her car is used to transport these functional requirements and sometimes even acts as a mobile studio. When we compare the creative practice of Jill and Faye we see a marked contrast. Faye does most of her painting away from her studio, whereas Jill rarely paints away from her studio. This suggests a fundamental difference in their inspirational needs.

6. CASE STUDY 3 – KERRY

Kerry is a photographer and photography teacher who has been practising his art for over 20 years. He now takes digital photographs. The darkroom is no longer used for processing photographs; digital photographs are "processed" using a Macintosh computer and Photoshop software. Although his medium has changed to digital format he is still inspired in much the same way as when photography was chemical-based. He is inspired by physical locations and also often uses human models in those landscapes. He is also inspired by the work of other photographers, often displayed on the Internet. Kerry's case does demonstrate that not all those working in digital media are inspired by sitting at the computer. His inspiration is generally outside, in nature, whereas the computer is mostly used for the functional processing of his images.

7. CASE STUDY 4 – THE VENERABLE BHIKKU SUMEDHA

The Venerable Bhikku Sumedha is a Buddhist monk and artist, who lives in a simple and remote cave near Kandy in Sri Lanka. The cave is used as a place for shelter, a retreat for meditation and also a studio for painting. Bhikku Sumedha creates his paintings in the middle of the night by candlelight. His inspiration for his paintings comes from his meditation. His paintings act as a form of communication with the outside world for this silent and reserved monk. Although his paintings are sold to eager Western buyers his motivation for conveying these images to the world is his desire to stimulate in the viewer certain thoughts on meditation and reflection. Bhikku Sumedha was not always a Buddhist monk; in his youth, he was a professional artist in Europe. His artistic style has changed significantly in that time, and whereas he used to paint as a career he now paints as a form of meditation.

8. DISCUSSION

From our previous studies we had noted a significant split in the creative process between initial inspiration (finding the problem) and creative development (solving the problem). This distinction was again evident in these studies. Our studies also suggest that these two different stages may require different technological support. Our current studies also indicated another distinction. Weintraub [10] suggests that there are two distinct groups of artists: those who are inspired from *within* and those that are inspired by *external* influences. The first group tend to meditate, take drugs, record their dreams or excavate their memories to find inspiration, whereas the second group tend to be more influenced by factors such as nature, politics, injustice etc. Our admittedly small sample of artists tends to confirm Weintraub’s analysis. Jill and Bhikku Sumedha are clear examples of internally inspired individuals. For them, their working creative environment, whether a studio or a cave is their inspirational environment. Jill tries to create the conditions for her memories and dreams to emerge and the Buddhist monk requires the right conditions for meditation as a basis for his painting. Faye and Toby are examples of externally inspired artists. For them it is nature or the human form in nature that inspires them and the studio is simply a place where their inspirations are brought to completion. The internal-external inspirational divide may relate to what Czikszentmihalyi [3] describes as the distinction between micro and macro creative environments. Microenvironments relate to the immediate setting in which the person works whereas macroenvironments include the social, cultural or institutional context. Of course, all people work in both a micro and macro environment, it is just that the dominance of the influence may relate to whether they are internally or externally inspired.

9. IMPLICATIONS FOR UBIQUITOUS INSPIRATIONAL ENVIRONMENTS

Our studies into supporting creativity in general and creative or inspirational environments in particular are in the early stages, however one or two insights have emerged. Clearly the environment does play a significant part in creative inspiration. It seems that both micro and macro environments need to be considered. The relative importance of which will be dependent on the individual dominance of internal or external inspirational factors. Technological (possibly ubiquitous) influences in the microenvironment could be the control of light, sound or random images. Technological support of a macroenvironment is expected to be a more difficult task. It is hard to augment the inspiration of nature, but those whom are inspired by politics, social injustice or similar factors may value the technological connection of their working environment with the outside world. One significant design problem we face is that the conditions or attributes that make an environment inspirational are very personal. Perhaps the most valuable potential role for technology, and in particular ubiquitous technology in our busy and crowded lives, is to transform ordinary spaces into personal inspirational spaces when required.

REFERENCES

[1] Candy, L. and Edmonds, E. (2002) Explorations in Art and Technology, Springer-Verlag, London, UK.

[2] Candy, L. and Hori, K. (2003) “Creativity and Cognition” Comes of Age: Towards a New Discipline, Interactions, July-August.

[3] Czikszentmihalyi, M. (1996) Creativity: Flow and the psychology of discovery and invention, HarperCollins, New York, USA.

[4] Eales, R. T. J. (2004) The Artist and the Computer: Understanding the creative user, Proceedings of OZCHI2004, Wollongong, Australia.

[5] Eales, R. T. J. (2005) Creativity in Action: Some implications for the design of Creativity Support Systems, Proceedings of CHINZ’05, Auckland, NZ.

[6] Florida, R. (2003) The Rise of the Creative Class, Pluto Press, North Melbourne, Australia.

[7] Mayer, R. E. (1999) Fifty years of creativity research, In Sternberg, R. J. (ed.) Handbook of Creativity, Cambridge UP, Cambridge, UK.

[8] Shneiderman, B. (2002a) Creativity support tools, Communications of the ACM, **45**:10, pp. 116-120.

[9] Weiser, M. (1991) The Computer of the 21st Century, Scientific American, vol. 265, pp. 94-104.

[10] Weintraub, L. (2003) Making Contemporary Art: How today’s artists think and work, Thames & Hudson, London, UK.

From Selective Indulgence to Engagement: Exploratory Studies on Photolurking

Haliyana Khalid, Alan Dix
Computing Department, Lancaster University, Lancaster, LA1 4YR, UK
h.khalid@lancaster.ac.uk, alan@hcibook.com
http://www.hcibook.com/alan/papers/HCI2006-indulgence

Based on three empirical studies of user experience in photologs, this paper introduces the concept of *selective indulgence* as part of user behaviour when lurking and sharing photographs. Displaying personal and domestic photographs in the public domain invites both acquaintances and unknown audiences. Despite the apparently limited communication offered by photologs, the studies reveal how participants allow themselves to spend time lurking on other people's photographs almost every day. The act of photolurking is usually done alone and involves selection from large numbers of photographs of every kind. This careful choice of what to look at, leading to intense emotion and engagement, is what we term selective indulgence. This study opens up new insights to inform the design of user experience in digital photo sharing. In particular the paper highlights the intensity of individual emotion, the phenomenon of using global photo sharing sites to share images with people physically very close, and subsequent gossiping and discussion of the photos not using the supplied commenting mechanisms.

User experience, indulgence, engagement, photo sharing

1. INDULGENCE IN PHOTOLOGS

Indulgence is a rather peculiar word to use to describe user behaviour when interacting with computers although some have used it. The word 'indulge' in the Oxford Dictionary means to allow ourselves to have or do something that we know we will enjoy [1]. In real life, we like to eat ice cream especially in hot weather because we know we'll enjoy it. Or we allow ourselves to be pampered in a shopping spree after a rough day at work. And let's just come back to the ice cream indulgence. Imagine yourself walking on a hot sunny day. You suddenly want to have an ice cream. You go to an ice cream parlour and pause ... which flavour will indulge you more? You pick a flavour that you think will give you the most pleasure. You eat and you feel happy. You are practising selective indulgence.

This scenario shows how we make decisions on what will make us happy, and how these decisions are based on our past experience. So can we associate the word 'indulge' with photo sharing? Is lurking among other people's photographs indulging? Our exploratory studies on user experience in photologs do suggest that this is the case.

The traditional way to share photos includes friends discussing their photos, a family sitting together flicking through a photo album and reminiscing their memories, and remote acquaintances sending each other photos by email. But with the advent of digital cameras and camera phones, people can now take as many photographs as they want, ranging from scenery to mundane day-to-day activities. And there is a need to store and to share these massive photograph collections easily [2]. From their research on blogging activity, Nardi et al. [2] found many bloggers have a strong desire to incorporate their photo collection in their blogs especially those who blog about their life. So photologs or photoblogs came into the picture, allowing users to order digital photos systematically, often in chronological order. A photolog constitutes one dimension of photoware [3], with people sending and sharing photos at different times and different places. Among the interesting features available in a photolog is a storytelling column in which people write about the picture, trackback, a feedback column, and some privacy elements. The term photolog is sometimes interchanged with photoblog, moblog, wireless blog or visual blog [4]. For the past two years, increasing numbers of photologs have existed on the Internet and feedback from users is overwhelming. Flickr is home for hundreds of thousands of photographs with a massive user base all over the world, Fotopages.com currently contains more than 50 000 photologs in their directory, from many countries.

Although many photolog applications exist on the Internet, and research on photo sharing has discussed the technical development of photo sharing applications, there is little information on user experience detailing what type of photos they share, what they do when they are on a photolog, and what makes them indulge and engage in lurking. Our research hopes to fill this gap.

To date we have done 3 exploratory studies on photologs beginning with photologgers' experience in photologs, then a quantitative study on the types of photos posted in photologs, and the latest on understanding photolurking. Our studies so far have given us some interesting insights, for example photoblogs have given birth to photography enthusiasts who capture many aesthetically self-posed photos and non-people photos such as scenery and food. They have created a favourite pastime for some: both photoblogging and photolurking; and a business for others. Photoblogs on the Internet have created a new online community that consists of photologgers and photolurkers from different countries, cultures and languages. Users are motivated to use this application because it allows mass photo sharing with remote acquaintances and the ability to socialize with them and others. Storytelling or photoblogging in the photoblog only allows for asynchronous interaction, thus extending the virtual communication to offline communication for feedback.

The photolog has also changed our belief and behaviour when collaborating with other people. From our studies, we found that the photoblog has become a virtual family album that portrays many ordinary family photographs, including group photos and events like birthdays, travelling and graduation ceremonies. What was personal, belonging to and treasured only by close families and acquaintances, has now become public viewing like a big billboard or a reality TV show. And to some people, seeing other people's photographs (strangers, celebrities, long lost friends, exes, and secret admirers) is fascinating and unconsciously it is an act to know ourselves [5]. In a photolog, one can freely look without the owner's presence and awareness.

In this paper, we will concentrate on users' selective indulgence in photolurking, which was discovered in our latest study. It is supported by findings from our previous two studies on photologs.

2. PREVIOUS RESEARCH

While it is a recent phenomenon, photo sharing has already stimulated much research. For example, Frohlich [3] discusses possible photo sharing applications to suit different locations and times, and introduces a system called ëaudiophotography' [6], in which photographs can combine with an audio track. Balabanovic et al. [7] developed a device to support storytelling in photo sharing and there are other research developments in photo sharing. However, whilst growing fast, research into photologs in particular is still in its infancy. Doring et al. [8] describe the phenomenon of moblogging on the Internet. Moblogging is a new concept introduced by Justin Hall and Adam Greenfield, in which people can update their moblog from camera phones. In their writing, Doring et al. describe general issues pertaining to moblogs including the usage and the architecture of the application. Van House et al [9] also provide findings on the usage of camera phones and what types of photos people like to share. Unlike Doring, Van House et al [9] constructed a web blog that allows numbers of participants to share their photographs.

3. METHODOLOGY

Six participants with mixed educational background, age and gender were chosen to participate in our latest study (study 3). Participants that were chosen either have a photolog or are familiar with the application. All of them have digital camera and camera phone. Participants were visited at their homes and interviewed in front of their desktop. The study was conducted in two sessions. In session 1 participants were asked to fill in a questionnaire to gauge their demographic background, Internet usage and photo sharing practice. Participants were also asked to list all photologs that they have visited. In session 2, interviews were conducted with a fixed and open-ended questionnaire. Participants were recorded as they browsed photologs during the interviews. In this study we seek to understand photolurking, where users view other people's photologs but post very few or no messages in the photolog. Our findings in this study are motivated and supported by our previous two studies on user experience in photolog. Briefly, the first study (study 1) focused on photologger experience using photologs. Here, we sought to understand the types of photos shared, what users do when sharing, their communication and their perception of the application. Another quantitative study to see types of photos posted was conducted in August 2005. A random 255 photologs were visited and their photos were observed (study 2). In total, 4883 photos were collected and they were categorized accordingly. This quantitative study strengthens our initial study on photologs. The depth of understanding that we have gained from these exploratory studies and the new insights to emerge from them, balance and justify the limited generality of our focused sample and methods.

4. PHOTOLOGGER AND PHOTOLURKER

From our constant observations on several photologs, and findings from our studies, photolog users consist of photologgers and photolurkers. These groups together have created the new online community of the photolog. Our emphasis on photolurking in this paper results from our findings in studies 1 and 3. In both studies, we found participants do more *photolurking* than *photologging*. A person who most often views a photolog is what we call a *photolurker*. The word photolurker is derived from 'lurker' which means an online user who posts very few or no messages in an online community such as interactive mailing lists and bulletin board systems [10]. Based on this, we have coined *photolurker* to mean a person who likes to view other people's photologs but posts very few or no messages in the photolog.

A *photologger* is a person who uploads their photographs in a photolog. Our participants are less frequent photologgers. They do photolurking more than photologging. They upload their photos occasionally based on events. Their photologs are intended for small audiences consisting of family members, friends and online acquaintances.

Our participants are also *frequent* photolurkers, visiting other people's photologs almost every day. They are often motivated or inspired by professional photographers or other people's photographs, which explains their frequent photolurking. In online photolog communities like Fotopages, a photologger is usually also a photolurker but a photolurker is not necessarily a photologger, although, in some cases, a photolurker later becomes a photologger.

5. INDULGENCE IN PHOTOLURKING

Eighty per cent of our participants visit a photolog everyday. Participants usually visit photologs during their free time and breaks. Photologs are also visited when they need to relax. One of the participants said:

> P4: "Before I have my own photolog, I always visited my friends' photologs. I browse my friends' photologs during my free time and when I'm bored with my work. Usually it takes about 5-10 minutes in one photolog. Then I click other friends and other links as well. Sometimes it did get carried away especially when there's interesting stories and photos."

Although the participants live near to each other, they prefer to view photographs alone in their room. Viewing photographs alone might appear boring, but it gives total freedom and concentration to the viewer. Concentration does affect indulgence.

> P3: "Although my friend just lives next door, they often asked me to upload my photos on photolog:
> P4: "I know I can view my friends' photos by going to her room, but we'll end up doing something else. When I view photos in the photolog, it's like I'm in my own world and always indulge in the photologgers' photos and their experience. I look very closely to her photos, and really wish I have the experience too"

Photo content plays an important role in determining user indulgence. Participants in study 3 like to look at 'people' photos. They also like to see holiday pictures that combine beautiful scenery and people. Wedding photographs also attract our participants.

> P2: "I like to browse 'A' wedding photolog. Browsing his photologs makes me dream about my own wedding; what theme colour I want to use, what dress I want to wear and where to get all the services."

Although a 'people photo' is considered as mundane by some viewers, there is a demand for it in both photolurking and photologging. Generally, most of the photos posted in photologs are of the 'people photo' type. Our studies 1 and 2 support this general observation. In study 1, of 470 photos observed in participants' photologs, 75% were 'people' photographs, and in our later quantitative experiment of the 4883 photos classified 61.5% are people photos. One interesting observation is that there are often self-posed photos or self-portraits posted in photologs. In our studies, participant 1 from study 3 described how she likes to see her own face in her friends' photologs.

> P1: "I like to see people 's photos. But I'll spend more time looking at my own face. And when my friends upload photos, I'll spot my face and take a closer look. You just can't do it with others presence..."

Since a photolog offers thousands of photographs, the photolurker unconsciously selects photologs that they want to see. Their preferences are based on their past experience photolurking. From their selective indulgence comes emotion. Lurking among photographs alone reveals emotions and these are usually natural, not expressed for others to see, as described by Buck:

> "When a user is alone, she has little pressure to present a proper image of herself to other people, thus any emotion expressed under these circumstances is likely to reflect a natural emotional/motivational state [11]."

During our interview, users described their emotions when lurking.

> P1: "It makes me happy to see those photographsÖ..It reminds me about everything that happens in the event"

6. DISCUSSION AND FUTURE WORK

This study has improved our knowledge and understanding of user experience in digital photo sharing. The birth of the photolog has sparked new phenomena. To some, it has become a virtual family album that portrays many ordinary family photographs, including group photos and events like birthdays, travelling and graduation ceremonies. However, what was personal, for family and friends only, has now become public for all the world see. There are some photologs that act as a gallery, portraying fine quality photographs for others to learn and envy. Cohen [12] describes how photobloggers like to take ëreal life' pictures; life as it happens, the small stuff, candid shots and anything that can tell the readers about the photoblogger's life or that she wanted to tell to others. The variety of real life photographs offered by photologgers invites many visitors, both acquaintances and strangers. Photologs are visited to keep update with friends' and strangers' lives, and keep up to date with the latest technology and photography skills.

This paper has provided some insights on photolurking and introduces the act as an indulgence for some people. But since a photolog site is a big public space with massive numbers of photographs of every kind, the users select what they want to see that will give them pleasure; so this interaction is a selective indulgence; and with it comes engagement and emotion. Although it is hard to justify explicitly, their ability to remember what they have seen and to describe emotionally, suggests they are deeply engaged in their experience of photolurking. Our participants chose to be alone to enjoy the photolurking experience. The freedom of being alone allows people to concentrate on their lurking, to be in their own world, to fantasize, to laugh and to remember. Most of them recall what they see and are able to discuss the photos later with their friends outside the photolog. They like both their own photographs and those of strangers. The fascination with images by and of others brings to mind Barthes' analysis of published photographs in *Camera Lucida* [13], which focuses on the *punctum*, the detail of very personal significance in a very public photograph.

For designers or builders and researchers on photo sharing applications and online communities, this introduction to photolurkers and their indulgence in photolurking provides new challenges for the management and support of online photo sharing applications and opportunities for future research. The two obvious uses for a photolog are to display photos for distant friends and also to show them publicly to unknown people. The latter is implicit from the public listings of photos and so in some ways photolurking is an expected phenomenon. However, this paper highlights unexpected aspects of photolurking: the intense emotional engagement of the indulgent experience, the sharing of photos with those near by through a global website, and the out-of-band discussions about photologs whilst not leaving online comments.

REFERENCES

[1] Pearsal. J, (2001) *The Concise Oxford Dictionary*, 10th Rev Edition. Oxford University Press. Oxford
[2] Nardi, B., Schiano, D. and Gumbrecht, M. (2004). Blogging as social activity, or, would you let 900 million people read your diary? In the *Proceedings of Computer Supported Cooperative Work*, Chicago, Illinois, USA, 2004, pp. 222ñ231, The ACM Digital Library (http://portal.acm.org)
[3] Frohlich, D., Kuchinsky, A., Pering, C., Don, A. and Ariss S. (2002). Requirements for Photoware. In the *Proceedings of Computer Supported Cooperative Work*, New Orleans, Louisiana, USA, 2002, pp.~166ñ175, The ACM Digital Library (http://portal.acm.org)
[4] en.wikipedia.org/wiki/Photoblog
[5] Lindgren, H. (1973). *An Introduction to Social Psychology*. Wiley, New York
[6] Frohlich D. (2004). *Audiophotography, bringing photos to life with sound*. Kluwer Publication, The Netherlands
[7] Balabanovi´c, M., Chu, L. and Wolff, G. (2000). Storytelling with digital photographs. In *Proceedings of CHI 2000*, pp. 564-571. The Hague, the Netherlands. The ACM Digital Library (http://portal.acm.org)
[8] Doring, N. and Gundolf, A (2005). Your life in snapshots: mobile weblogs (moblogs), In: Glotz, P, Dertschi, S, Locke, C (Eds), *Thumb Culture*, Transaction Publishers, New Brunswick (USA) and London (UK)
[9] Van House, N., Davis, M., Ames, M., Finn, M. and Viswanathan, V (2005). Late breaking results: posters: The uses of personal networked digital imaging: an empirical study of cameraphone photos and sharing. In *Proceedings of CHI '05*, Vol 2, pp. 1853ñ1856. Portland, OR, USA. The ACM Digital Library (http://portal.acm.org)
[10] Preece, J. (2000). *Online communities : designing usability, supporting sociability*. John Wiley, New York
[11] Buck, R. (1976). *Human motivation and emotion*. John Wiley, New York
[12] Cohen, K., (2005) What does the photoblog want? *Media, Culture and Society*, 2(6), 883-901
[13] Barthes, R. (1993) *Camera Lucida* (trans. R. Howard). Vintage, London.

Rules of Engagement: Design Attributes for Social Interactions

Paul André, Martin Chapman, Alex Frazer, Charles Hargood, Alex Hayton, Clare Hooper, Gavin Willingham, Kirk Martinez, m.c. schraefel[1]
Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, UK
[1]*mc+6012@ecs.soton.ac.uk*

We present a taxonomy for the design of workplace "break" spaces. The taxonomy can be used to identify aspects of current spaces that are either successful or problematic. From this analysis, we demonstrate how the taxonomy can be used to identify opportunities for computer mediated augmentation of spaces, and how such designs can be validated against this taxonomy.

Taxonomy, communal spaces, design review, social interaction, social space, engagement

1. INTRODUCTION

In this paper, we consider the design attributes necessary for computer-mediated support of effective social-presence interactions between work activities. This interest has been motivated by a fire which destroyed our School's coffee room, a key social space for over 300[1] academics, support staff, postgraduate students and researchers who used it daily. The coffee room was a well-used place for a variety of social interactions, from impromptu meetings with colleagues from different groups, to simply seeing who is about while getting a coffee. Because of this loss, we have had a unique opportunity to look comparatively at the specific attributes for effective socialisation in spaces designated to support these activities. Since the coffee room burnt down, our entrance foyer, recently refurbished, has been used as one substitute, but has not been nearly as successful as a social space. Likewise, hallway kitchenettes have been utilised by many people, but mostly for quickly grabbing a coffee, not for taking breaks with others. The goal of our work has been two-fold: first, to investigate the specific differences between these three places to understand why one worked well (many people participated) and the others less well (many people who used the old space do not use the new spaces; nor have they invented alternative/replacement practices), and second, since we cannot in the near term generate new physical space, to see how these differences may be addressed by digital rather than physical solutions. Our approach is informed by Dix's Christmas Crackers work [1], in which he "deconstructs" both the physical and affective properties of Christmas crackers in order to see how these attributes might be translated from a physical to a digital experience. By carrying out such an analysis of the coffee room, in combination with ethnographic studies of the alternative spaces, we developed a taxonomy of practical and affective attributes such spaces seem to need to support to be successful. We then use this taxonomy to propose three candidate digital artefacts: KitchenSync, tableTOP and EC-Chess, to reintroduce some of the practical and affective affordances that have been lost.

2. RELATED WORK

Herbsleb *et al.* examine the role of contextual awareness [2], informal communication [3] and tools such as instant messaging (IM) [4] for distributed work. Their analysis shows that co-location is critical to teamwork. Informal communications are thought to be highly effective in the work environment, yet tools such as IM tend to be rejected [4]: many workers associate typing at a computer, for communication or other purposes, with work, yet perceive informal face-to-face communications as desirable. Indeed, it has been suggested that 25% - 70% of people's work time is spent in face-to-face interaction (variations due to job type) [5]. It has also been suggested that chance meetings trigger useful conversation [6]: other works have tried to simulate this over a distance [7, 8].

A number of systems with the purpose of indicating presence are proposed by Hindus *et al.*, including *InTouch*, and the *Intentional Presence Lamp* [9]. These systems use combinations of light, sound and imagery to indicate one user's presence to another. Other systems [10] have used different indicators, such as avatars, names, or photos. Greenberg and Rounding present *The Notification Collage* [11], supporting conversation through the use of "Sticky Notes". The universal visibility of these notes on communal displays encourages new users to join in, after "overhearing" publicly visible conversations. The notes' persistence

[1] http://www.ecs.soton.ac.uk/news/fast-facts.php

allows asynchronous conversations to take place between users. However, as the collage elements are universally accessible, it was possible for them to be moved or hidden entirely by new elements, causing users to “miss” their messages. Privacy issues were also uncovered with the system's video capture functionality, where users could potentially be filmed against their knowledge. Although some interesting work has been done in supporting social interactions [12], less work has been done on supporting the social engagement between work activities, such as during coffee breaks. One approach, using video for co-presence across group kitchens [13], was largely a failure due to privacy concerns. Privacy was also a concern in Fish *et al*.'s *VideoWindow* system [14] and the *Montage* system [15] by Tang *et al*.

3. METHODOLOGY

With the aim of comparing and contrasting the coffee room and substitute spaces, we set out to identify, in each space, the activities and use of the space, and how they are afforded. We undertook this elicitation by conducting a series of observations of the foyer and interviews about all three spaces.

3.1 Observations

Based on local knowledge and preliminary observations, we identified the main stakeholders for the coffee room: lecturers, researchers, postgraduates, staff, and undergraduates. We focused our observations on how and when these groups made use of the foyer. Three investigators observed the foyer at various times over the course of several weeks, totalling 14 half-hour observation periods. To determine times for observation, we made use of existing presence data. This presence information correlated with knowledge of designated activities that were scheduled during the week, such as student interview times, seminars and transitions between classes.

3.2 Interviews

We conducted twenty-four structured interviews across stakeholder groups. The consensus was that, to those that used it, the coffee room was missed. The reasons varied depending on the stakeholders’ use of the space: lecturers missed the opportunity to socialise with postgraduates and researchers; postgraduates missed the change of atmosphere afforded by a dedicated coffee space; undergraduates, who previously used the space for project meetings, missed the availability of such a venue. All parties except undergraduates, possibly due to their less frequent use of the space, felt that chance meetings and awareness of other department members were important activities that had been lost. In addition to what was missed from the coffee room, negative opinions of the foyer were also highlighted: the coffee available from the foyer was unanimously derided and senior stakeholders - particularly lecturers - felt that the foyer was not a private enough space to discuss certain topics.

4. ANALYSIS OF RESULTS

From analysis of the interviews and observations, we derived seven categories that recurred across participant communities. In Table 1, we present a preliminary taxonomy of (1) the values, and (2) three examples of how these values map to the three physical spaces: the original coffee room, the foyer and hallway kitchenettes. By mapping these spaces to the taxonomy, the table demonstrates how these values may enable a way to interpret why one physical design works better than another, and also highlight design opportunities for where virtual intervention may be able to address reduced values. Values are grouped into two categories: artefact and activity. Artefacts are attributes of the space itself, while activities are interactions supported by the space. The degree to which the space supports these values is represented in the colour coding: strong (green), mediocre (yellow) or poor (red).

4.1 Description of values

Artefact values. A **lure** can be a compound artefact such as the presence of both *good* coffee and colleagues. **Environment** is the design of the break space. **Awareness of others** is a boundary value between an artefact and an activity: while presence is a mental rather than physical artefact, to determine presence, one has to act.

Activity values. The activity of **breaking away from work**, which means changing location and task, was a recurrent value that emerged in interviews. **Serendipitous meetings, semi-planned meetings** and **socialising** were all seen as critical activities to be supported for effective coffee break interactions.

5. DESIGN METHOD

In the following section we present three examples of using our taxonomy with Dix’s Christmas Cracker design method [1] to map affective attributes and physical affordances of an actual artefact to its virtual counterpart. In our case, rather than design a replacement artefact, we present supplemental ones to address missing physical values with virtual artefacts designed to add these values to the space.

5.1 KitchenSync – Values addressed: awareness, engagement, lure
An interactive screen is placed in each kitchenette, which shows who is currently using, and who has used the kitchenette in the last ten minutes. People can monitor this activity from their office, or from other kitchenettes. Presence may potentially be detected by attaching cheap RFID tags to a mug, which also allows people to opt in or out of the system by their choice of mug.

	Value	**Affordance**		
		Coffee room	**Foyer**	**Kitchenette**
Artefact	**Lure (enticing factor)**	Good coffee and contact with colleagues, proximity	Poor quality coffee	Good coffee, microwave, proximity
	Environment	Enclosed, purpose designed, social space, windows, multiple tables	Clinical, waiting room feel. Office work (reception) nearby. Transient.	Small, no windows, just a place to make coffee
Activity	**Awareness of others, presence**	Achieved by looking round space, or asking who has been here	Good awareness, but too transient	Too small to socialise in, can't walk through it
	Taking a break (change of location & activity)	Away from offices, different setup	Too professional, feels like still in work	Limited space; standing room only
	Engagement (unplanned)	Forced to walk past tables to coffee, through-traffic	People pass through, but often coming to/leaving work	Too small to socialise in, no through-traffic
	Semi-planned meetings	Table arrangement provides focus. Cannot book the room.	Area is often empty, perceived space for only one group	Limited space for meetings
	Socialising	Many tables, suitable for different groups to meet	Perceived space for only one group	Limited space for socialising

TABLE 1: Mapping values against affordances of social spaces

Awareness - The system provides lightweight, non-intrusive awareness of others' presence. *Semi-planned meetings* are afforded by the ability to choose to respond to people's presence by physically going to the kitchenette to meet. *Serendipitous engagement* is enabled by persons noticing the presence of a colleague in another kitchenette, and signalling their interest in meeting the colleague through a lightweight mechanism such as touching the colleague's avatar on the screen. This last feature might only be available to kitchenette users. By offering this asynchronous communication, initiated virtually, the system facilitates the kind of chance physical meetings previously valued.

5.2 tableTOP (tableToOccupyPeople) / EC-Chess –
Values addressed: awareness, engagement, environment, break, socialising, lure
An interactive table-top system is proposed for the foyer, offering virtual awareness from the KitchenSync system, as well as virtual postcards for others to read, perhaps leaving their topics of discussion for viewing or for adding their own comments. The table also provides games such as chess, sudoku, or Go.

Awareness. Due to the transient nature of the foyer space, people are rarely present for very long, making it unlikely that they will remember others who were previously in the space. By recording this information in the table and relaying it back, the levels of awareness afforded by the more social situation in the coffee room is recaptured.

Engagement. A person can leave postcards for others, which are shown automatically when this person's presence is detected (possibly by their RFID mug). The recipient can leave subsequent messages in reply, providing an opportunity for asynchronous conversation, or use the notes to share information with anyone who sits at the table. The games can be offered in an 'open' style, where anyone present can take the next move, regardless of whose game turn it is. By detecting who takes each move, a play history can be displayed, encouraging users to ask others why the moves were made, or offer suggestions for alternative future strategies, aiming to recapture the *serendipitous conversations* found in the coffee room. The system could also be used to record private games. The number of wins and losses could be recorded, pairing up users of similar skill in future matches. Again, this could encourage spontaneous conversation between people, re-affording the kind of spontaneous engagement currently missing from the building's interim recreational space.

Environment / Break from work. Our application of the taxonomy shows the foyer is an unwelcoming space. We postulate that by providing awareness, engagement by messages and games we will improve the environment (*socialising* it by introducing ludic [16] qualities), but also potentially improving the *lure*, where the system itself becomes a reason to go to the foyer and hang out.

5.3 Generalisation

We have shown that the values in the taxonomy are useful both for understanding the strengths and weaknesses of a space, and for designing solutions to address the weaknesses and enhance the strengths. Our application of the taxonomy demonstrates its generalisability for assessing coffee break spaces.

6. CONCLUSIONS AND FUTURE WORK

By comparing and contrasting the loss of an effective social space with substitutes we have developed a taxonomy of seven key values for break spaces. The taxonomy provides a framework for the analysis of physical spaces in terms of their affordances for social interaction. This analysis highlights strengths and weaknesses which can be used to inform design requirements. We have demonstrated how this assessment can be used in the design of digital systems to augment these spaces to better support the rich social activities that occur during coffee breaks.

We are currently prototyping the design ideas presented. We are interested in looking at how many identified weaknesses need to be addressed, or strengths augmented, in order for a proposed virtual system to improve human engagement in a break space. We are also interested in testing how the taxonomy may translate, or need to be extended, to support analysis and design of other spaces for social interaction, from pubs to parks.

REFERENCES

[1] Dix, A. (2000) Deconstructing Experience - pulling crackers apart. In *Funology: From Usability to Enjoyment*. M. Blythe, K. Overbeeke, A. Monk and P. Wright (eds.) Dordrecht, the Netherlands: Kluwer.
[2] Espinosa, J., Slaughter, S., Herbsleb, J. (2002), Shared Mental Models, Familiarity and Coordination: A Multi-Method Study of Distributed Software Teams, *23rd International Conference on Information Systems* 425 – 433, Barcelona.
[3] Herbsleb J., Grinter R., (1999), Splitting the Organization and Integrating the Code: Conway's Law Revisited, 85-95, In *Proceedings of ICSE 1999* LA, CA.
[4] Herbsleb, J., Mockus, A., Finholt, T., & Grinter, R., (2001), An empirical study of global software development: distance and speed, *International Conference on Software Engineering*, 81 – 90, Toronto.
[5] Whittaker, S., Frohlich, D., & Daly-Jones, O.,(1994), Informal workplace communication: What is it like and how might we support it? In *Proceedings of CHI '94*, 130-137, ACM Press.
[6] Jeffrey P., (2000), Forum Contact Space: Serendipity in the Workspace, 331 - 332, In *Proc. CHI 2000*, ACM.
[7] Isaacs, E.A., Tang, J.C., & Morris, T. (1996), Piazza: A desktop environment supporting impromptu and planned interactions. In *Proceedings of CHI '96*, 315;-324.
[8] Zhao Q. A., Stasko J. T., (2000), What's Happening? The Community Awareness Application, *Proceedings of CHI 2000* 254 - 254, ACM.
[9] Hindus, D. et al., (2001), Casablanca: Designing Social Communication Devices for the Home. *Proc. SIGCHI Conference on Human factors in Computing Systems*.
[10] Leigh, J., Johnson, A.E., (1996), Supporting Transcontinental Collaborative Work in Persistent Virtual Environments. In *Computer Graphics and Applications*, IEEE.
[11] Greenberg, S., Rounding, M., (2001), The Notification Collage: Posting Information to Public and Personal Displays. *Proc. ACM Conference on Human Factors in Computing Systems*.

[12] O'Hara, K., Perry, M., Churchill, E. and Russell, D. (eds.). (2003) *Public and Situated Displays - Social and Interactional Aspects of Shared Display Technologies*. Kluwer Academic Publisher, London.
[13] Jancke, G. et al., (2001), Linking Public Spaces: Technical and Social Issues. In *Proceedings of CHI 2001*, Vol. 3, Issue 1
[14] Fish, R. S. et al., (1990), The VideoWindow System in Informal Communication. *Proc. ACM Conference on Computer-Supported Cooperative Work*.
[15] Tang, J. C. et al., (1994), Supporting Distributed Groups with a Montage of Lightweight Interactions. *Proc. ACM Conference on Computer-Supported Cooperative Work*.
[16] Gaver, W. (2002) "Designing for Homo Ludens," in *i3 Magazine*.

Strengthening Community with Embodied Social Networks

Tom Hope, Masahiro Hamasaki, Yutaka Matsuo, Takuichi Nishimura
Information Technology Research Institute, 2-4-6 Aomi, Koto-ku, Tokyo 135-0064, Japan
tom-hope@aist.go.jp

Social Network Systems have become highly applicable to everyday life, but continue to remain on the desktop for most users. This paper reports on initial analysis of groups interacting with a social network system in the real-world, in this case a conference setting. The system, Polyphonet Conference, and its RFID card interface together allows rich interaction between users. Ethnographic observation of user interaction, with the use of video data collected at the time of use, was used to assess what social benefits may be afforded by the system. This paper suggests that the act of adding to one's network may in itself help to generate and strengthen community. Using the notion of folk computing, community is seen to be generated particularly well when it occurs via the embodied action afforded by a combination of virtual web-mining and subsequent user authorship.

Community, Polyphonet, RFID, folk-computing, social network.

1. INTRODUCTION

This paper explores the meeting of two lines of research: the development of social network mining and visualisation, and human-computer interaction in temporary communal spaces. In recent years research has begun to move away from using social networks as simply a unit of analysis to a more proactive use of networks to connect the interests of individuals and groups [1, 2]. Social matching software aims to bring together individuals who have similar interests or characteristics [3] and may include the ability for users to suggest elements of their own identities to be used in the analysis and subsequent matching. This trend has also provided opportunities for the development of systems in which users directly view and alter their social networks in virtual space, including many popular online social network systems [4,5,6]. In this paper we present an initial analysis exploring the connectivity that occurs around and through social network technology in a real-world setting. We are interested in how humans interact with a social network system in a semi-public space and as co-present, rather than distributed, communities [7]. Though there has been continuing research into collaborative work in face-to-face settings, notably in the form of studies of technology use in the work place [8,9], it is only relatively recently that attention has turned to the community building that occurs in other settings.

Inspired by the vision of folk-computing [7], in which communities are strengthened through their use of oral and tangible 'folk-games', and work on embodied interaction in HCI [10], the following analysis illustrates how an interface and social network system may help build real-world communities.

2. POLYPHONET SYSTEM DESIGN

Polyphonet Conference has been developed with the aim of enabling participants of conferences and other events to build relationships with others present, based on their previous connections via other colleagues, potentially with future impacts on research collaborations. The system consists of two aspects, an online web-browser in which users can see their social network consisting of participants of the conference, and a kiosk-based interface, where users can log in to the system, an act that additionally affects the network. Prior to the event, the World Wide Web is crawled using registered participants to the conference as the 'nodes' of a network. The web-mining then uses a search engine to establish the relationships (if any) between participants (represented as 'edges' on a network) and the strength of those relationships, according to the number of hits on a search engine (the process is described more fully in [11]). The result is that a social network consisting of the participants of the conference and the strength of their relationships can be displayed to users.

Figure 1 shows a screenshot of the page as seen by a user once they log into the system, whether at their own computer or a kiosk. This homepage, known as 'My Page', provides the user with information about members in their network, shown on the middle-left of the screenshot. To the right is a list of members with strong relationships in the network. Additionally, information is given about papers or poster sessions that the

user has bookmarked. To the bottom of Figure 1 can be seen the user's social network diagram. This can be expanded to full-screen dimensions by the user.
In addition to the functions provided by the online browser, on-site functionality is given in the form of a kiosk-based interface, where up to three users are able to log in together, thereby viewing their joint network diagram.

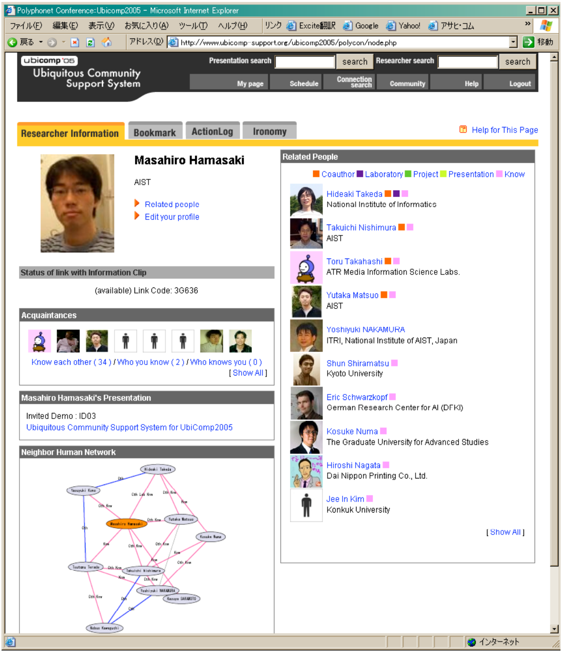


FIGURE 1. Polyphonet 'My Page'

The social network is built prior to the conference if users have registered, or at the conference, in which case some time is needed to mine the Web. Users are then able to modify their network, by adding other conference participants through two additional methods. Firstly, they can select a participant's name and add them to the network by selecting a 'know' link. These types of links are then marked as 'know-links' on the diagram. Secondly, they can add members and be added to other members networks by logging into the system at a kiosk together. At the conference each member is given an RFID card, which is then registered to his or her name. This allows them to log into the system at kiosks individually—in which case they are taken to their My Page—or together, in which case the system automatically adds members to each others' networks (if they are already not part of the network). This latter method also labels the network links (edges) as 'touch links'. Users are thus able to trace the development of the network from its initial web-mined form, to the diagram as modified at the actual event. Figure 2 provides a screenshot of a joint network as seen when logged in together with the RFID cards. The three users images are at the left, coloured according to the RFID reader that they used to log in.

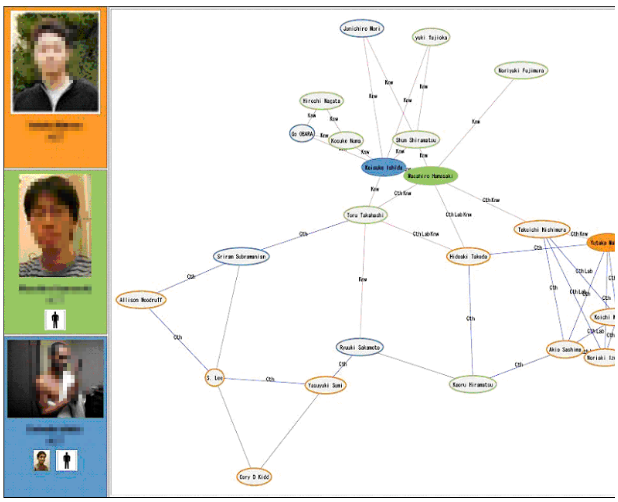


FIGURE 2. Polyphonet joint network of three users

3. USER STUDY

3.1 Methodology

Polyphonet was tested at two conferences in 2005, *JSAI 2005* and *Ubicomp 2005*, each held inside Japan and having close to 500 participants over a period of three days. In order to observe the usage of the system, video and audio data of users at a kiosk was gathered over the entire length of the conferences from multiple angles. This, in addition to the access log data on the database and ethnographic observation by the project's researchers at the time, has allowed detailed and repeated viewing of users' actions. There is now a strong discipline of analysis of conversation around and with technology [12]. Incorporating ethnomethodological study [13] with recent perspectives of technology as equal actors [14] is providing a viewpoint of the interaction with the system that incorporates both the affordances of the technology and those of the human users. We present our initial results below, though further development and analysis of Polyphonet continues.

3.2 Results: Embodied Social Networks

While it was not possible to view directly the users interacting with the system on their own computers, viewing the video of their interaction at kiosks has proved to be enlightening. Two observations are pertinent to the aim of Polyphonet as a community-building tool. Firstly, once introduced to the system, users were keen to add other members to their networks. They added via 'know link' those conference participants who they felt they had research relationships with and also those that they had no formal relationship with, but whom they knew from other conferences. This illustrated the difference between the system as envisaged by web-mining and that as understood by many users. Most members' online identities (where they use their full name) may be expected to be research or work-related. The web-mining will therefore build work-related networks from this data. Yet, users desired to modify these networks to more accurately reflect their own perceived relationships, which may be personal or friendship-based.

Secondly, users would discuss quite intensely the meanings of the networks in relation to their communities as perceived personally. When users would disagree with the initial network as provided by the web-mining, their negotiation involved the construction of a modified network and then passing this orally to the other members when logged in or standing by the kiosk monitor together. Members would deliberately log in together with their RFID cards to add each other their networks, including serendipitous meetings of acquaintances from previous events. While web-mining provided the base network, conference participants used the embodied nature of the RFID cards and kiosks to modify the social network visualisations; the embodied action being one of both technologically building a network and socially building relationships and community.

This latter observation points to the role that learning took place in the use of Polyphonet at the conferences. Individuals who had adjusted their social networks on the system acted on the boundary between new users and Polyphonet, utilising the tangible interface of RFID, kiosk and network, to introduce new users to the system via 'touch link'. They thereby acted as 'brokers' [15] to their own communities, other members of which were often with them physically at the time.

4. DISCUSSION

The results of our deployment of Polyphonet to date provide an insight into embodied computing [10] in relationship to temporary co-present communities. In their work on Folk Computing, Borovoy et al [7] suggest that three features are important to develop community with technology. These are authoring, passing and tracking. Polyphonet has two of these features. Principally, the creation of know-links and touch-links by users is an act of authoring. In their papers on i-balls and memetags by Borovoy et al [7,16], subjects created small digital items that could be passed from user to user. In Polyphonet, users authored the social network itself via their modifications online and with RFID. Folksonomy [17]—the act of tagging often used in online social network systems—may be regarded as another act of authoring. We suggest that, seen in this light, the use of embodied social network systems can be a form of contemporary 'folk game'.

Action around a kiosk with multiple users is both a physical and an oral form of 'passing'. Users orally negotiate and explain their networks to each other, and, they are able to then physically modify their networks accordingly. Unlike i-balls or memetags, which may be primarily authored individually, social networks exist jointly between users; they consequently suggest actions of passing networks *within* the negotiated process of authoring them. In order for users to play this folk game, it is necessary to provide some form of resource with which to play. In Polyphonet, the web-mining provided such a basis. Using a tangible interface, such as the RFID cards used with Polyphonet, appears to provide another resource; in a sense this being another 'actor' that prompts interaction and hence the basis for community. The manner in which users could easily log in with other members and then discuss the modifications done or deemed necessary, was learned and then taught to others. Folk games also take this form.

Though the system enabled authoring and passing, in order to more fully explore the folk computing aspect of Polyphonet, it may be necessary to allow more tangible passing (such as the creation of mobile devices or tags) and subsequent tracking of authored networks. The passing we observed took place mainly around the kiosks. Though the fixed nature of these kiosks afforded opportunities for interaction, users may benefit from the ability to carry their networks away from the location to other spaces and settings. Additionally, greater tracking of the development of those networks over the course of events and afterwards should provide another resource for users to explore their own community. One of the most significant elements of community-building can be the intimacy gained thorugh a sense (real or imagined) of a 'shared' history [18,19]. Giving users the ability to view a history of the development of their social networks may provide this resource.

5. ACKNOWLEDGEMENTS

This research has been supported by NEDO (New Energy and Industrial Technology Development Organization) as the project ID of 04A11502a.

REFERENCES

[1] Mori, J., Ishizuka, M., Sugiyama, T. and Matsuo, Y. (2005) Real-world oriented information sharing using social networks. Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work, Sanibel Island, Florida, pp.81-84. ACM Press, NY.
[2] J. Goecks and E. D. Mynatt. Leveraging social networks for information sharing. In *Proc. ACM CSCW 2004*, pp. 328–331, 2004.
[3] Terveen, L. and McDonald D.W. (2005) Social Matching: A Framework and research agenda. ACM Transactions on Computer-Human Interaction. pp.401-434. ACM Press, NY.
[4] Orkut. (http://www.orkut.com)
[5] Friendster (http://www.friendster.com)
[6] Mixi (http://www.mixi.jp)
[7] Borovoy, R., Silverman, B., Gorton, T., Notowidigdo, M., Knep, B., Resnick, M., and Klann, J. (2001) Folk computing: revisiting oral tradition as a scaffold for co-present communities. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Seattle, Washington, United States. CHI '01. pp.466-473.ACM Press, New York, NY.
[8] Suchman, L. A. (1987) Plans and Situated Actions: the Problem of Human-Machine Communication. Cambridge University Press, Cambridge.
[9] Ueno, N. (2000) Ecologies of Inscription: Technologies of Making the Social Organization of Work and the Mass Production of Machine Parts Visible in Collaborative Activity. Mind, Culture & Activity. 7 (1&2), 59-80.
[10] Dourish, P (2001) Where the Action Is: the Foundations of Embodied Interaction. MIT Press, Cambridge.
[11] Matsuo, Y., Mori, J., Hamasaki, M., Ishida, K., Nishimura, T., Takeda, H., Hasida, and Ishizuka, M (2006): POLYPHONET: An Advanced Social Network Extraction System. Proceedings of WWW2006, 23-26 May.
[12] Hutchby, I. (2001) Conversation and Technology: From the Telephone to the Internet. Polity Press, Cambridge.
[13] Garfinkel, H. (1996) Studies in Ethnomethodology. Polity Press, Cambridge.
[14] Latour, B (2005) Reassembling the Social: An Introduction to Actor-Network-Theory, Oxford University Press, Oxford.
[15] Wenger, E. (1998) Communities of Practice: Learning Meaning and Identity, Cambridge University Press, Cambridge.
[16] Borovoy, R., Martin, F., Vemuri, S., Resnick, M., Silverman, B., and Hancock, C. (1998) Meme tags and community mirrors: moving from conferences to collaboration. Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work , Seattle, Washington, November 14 - 18. CSCW '98. pp.159-168. ACM Press, New York.
[17] Millen, D. R., Feinberg, J., and Kerr, B. (2006) Dogear: Social bookmarking in the enterprise. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Montréal, Québec, Canada, April 22 - 27. CHI '06. ACM Press, New York.
[18] Tönnies, F (1955 [1887]) Community and Association, Trans. Loomis, P, Routledge and Kegan Paul, London.
[19] Vetere, F., Gibbs, M., Kjeldskov, J., Howard, S., Mueller, F., Pedell, S., Mecoles, K., and Bunyan, M. (2005) Mediating Intimacy: Designing Technologies to Support Strong-Tie Relationships. ACM CHI 2005, Portland, Oregon, USA

HCI Research in the Home: Lessons for Empirical Research and Technology Development

Dorothy Rachovides and Mark Perry
School of Information Systems, Computing and Mathematics,
Brunel University, Kingston Lane, Uxbridge, Middlesex, UB8 3PH, UK
{dorothy.rachovides, mark.perry}@brunel.ac.uk

Information and communication technologies have begun to permeate our home environments under the auspices of the ubiquitous and mobile computing and information appliance movements. Yet the home is a very different environment to the workplace which has been the focus of the majority of HCI research. We have relatively few studies of information behaviour within the home or of data gathering methods that will allow researchers to investigate current domestic information and communication practices and needs. In parallel with this, there are a number of issues arising from designing technology for the home environment that are specific to the particularities of home life. Our experience in such a study of data collection and technological intervention within the home environment has provided detailed insights into these issues and problems, and we report on these here, presenting suggestions for future research programmes within the home.

Shared displays, home life, communication, user studies, interaction design, multimodal.

1. INTRODUCTION

The on_message@home project is investigating household communication to support the design of a home-based messaging system, and to do this we have been examining communication practices and information sharing between members of the home. We use the term ‘messaging’ in its broadest sense, to include notes, voice messages, reminders, to-do lists and photographs that have been placed for viewing by others, as well as more formal communications, much as the family refrigerator and other notice boards or corkboards are used for. The implications of this research have led to the development of a prototype messaging system through which users can remotely ‘post’ messages to situated displays (situated, because display content is location specific) in their homes. We are prototyping a heterogeneous device environment for message posting, with mobile telephones sending SMS (text) and MMS (multimedia) messages, mobile devices connecting over Bluetooth or Wireless LAN (for example, to send photographs or mp3 files), remote web access via a PC over the Internet (to send notes, post interesting documents or other files), and by ‘posting’ messages through local and remote voice-based media (through a dedicated voicemail account). These multimedia messages can then be viewed on wall-mounted, interactive, displays that allow family members to view, create, retrieve, sort, discard, move and repurpose the material on them. This ‘ecology’ of appliances provides a complex platform for use and design, and we have focused our efforts on domestic messaging behaviour and patterns of use around this technology set.

A number of technical concerns arise out of designing such a system, but our primary concern has been on user interaction with the device set: this has taken two forms. 1) To develop interaction designs and principles for interface design to ensure the technology supports its users’ needs and that it can be used without becoming a demanding cognitive task. 2) To evaluate interaction around the set of devices, for example, enquiring if, why and how the technology changes family roles and relationships, and alters the balance of power within the family; how it may become a focus for certain types of information; how the device is appropriated for playing games and in jokes; whether it improves (or not) family event co-ordination; and how it affects the previously tacit monitoring and policing activities that family members may engage in. So what makes the home so different from the workplace? It has been argued that the methods that we use to communicate in the home are practically no different to those of the workplace. The resources and mechanisms (both social and physical) that we have available to us at work to interpret and act through are the very same ones that we employ at home. What differs is the context of use: the activities that users engage in and the household relationships create a related, but distinctly different, set of requirements for home-based IT systems.

2. STUDYING IT IN THE HOME ENVIRONMENT

In this section, we cover what we have found to be the important issues in studying and understanding a particular aspect of IT in the home environment (i.e. a multi-modal messaging system). These are many and diverse, and cover technical, interactional and social problems. When developing for the home, we have faced particular problems in making use of the existing devices in a networked system. Specifically, we have faced problems with getting the complex technology to interact with one another: the appliances that we use are often early on in their design lifecycles (for e.g., a GSM terminal/mobile telephone, PDA or Bluetooth), and development software (e.g. powerful and flexible APIs and SDKs) is simply not available, or requires substantial programming effort in order to perform simple actions that are not a part of the basic set of actions expected by their developers. The integration of multiple consumer devices onto computers, has also occasionally resulted in software/driver incompatibilities, and some consumer technology manufacturers have actually prohibited certain uses of their devices, denying us the opportunity to use them flexibly – a problem compounded by not putting this information in their pre or post-sale documentation. Other devices and software packages are not possible to control through high-level prototyping software, such as Macromedia Director. For a small team of interaction designers and social scientists, this is not a trivial problem. An example of this has been in integrating MMS and voicemail via a GSM Terminal. Of course, we can simulate this, but the purpose of the prototype is to investigate real world activities, and to develop any real understanding of practice, we need to have a semblance of the system's actual functionality.

Yet to place the problematic technological issues at the core of system development would be a mistake. The key concerns that we face in developing an effective design, as might be expected from an interaction-oriented project, centre around social issues. This is not to say that they are 'problems' as such, but they are aspects that may be problematic for the household to resolve, and may impact on the acceptance of the technology in the long run (or even rejecting an initial deployment of the prototype). As we have implied earlier, just because a technology is simple and does not greatly alter the functional activities of the household it can still have the potential for social disruption. This could occur through subtle changes to family roles (e.g. by distributing the role of the information gatekeeper) and relationships (e.g. enhancing displays of affection by allowing remote others to interact with those at home, and enriching the methods they can use). Some of these changes have the potential to shift the balance of power within the family or household. There is also a potential that explicit and publicly visible representation of communication could affect the previously tacit information monitoring and policing activities that adult or controlling family members may engage in. This is not to say that any of these resulting behavioural changes is negative, or that they are undesirable result of the study – we are as interested in understanding the transformative effects of networked communications technology at home as we are in developing useful and usable designs. By transforming activity, technology can provide fascinating insights about how family relationships operate, much as an ethnomethodological breaching experiment [1] can tell us about the maintenance of social order by breaching the commonly-held-to-be-true 'rules' of the home.

Another matter of concern to the households being investigated, and this may relate to any commercialised design, lies in privacy, and access to content. It is perhaps of greater concern to households for the prototype, who have opened access to their information by researchers. This is partly that they may have issues of embarrassment in opening up their private worlds, but also because this technology may allow access to sensitive information that could be useful to burglars or nefarious others (e.g. phone numbers, children's photographs, calendar information about when they might be on holiday, information that might be used against them in court). Even the integration of a web camera in the system to record video messages has given rise to questions about who might be able to access a video stream remotely (thieves, paedophiles or stalkers). These are serious practical issues to resolve if we are to ensure this technology has any hope of being installed. It must be remembered in these circumstances that household users may not be fully aware of the potential of the technology (which might be very limited in practice) in a way that technology developers would scoff at. Yet to get such a technology set adopted into the home will require these concerns to be addressed. Indeed, some of these concerns may be grounded in a very real danger, and they are not ones that we should attempt to explain away or deny.

3. EMPIRICAL STUDIES IN THE HOME

Until recently there has been little interest within computing and technology studies [2] in the area of domestic computing, but a number of studies of home life with an orientation towards Ubiquitous Computing, HCI and CSCW have begun to produce findings which can provide designers with an insight into the potential for domestic and leisure technology. These studies cover a broad spectrum of everyday domestic life, from mediating intimacy [3] to 'mothers work' [4], and calendar use [5] to the organisation and use of paper mail [6]; [7]. Within this growing research interest into the home, there is a thread of interest in which has begun to point towards the roles that different display surfaces play in the home, and we have extended this to look at the role of messaging in display surfaces. To do this, we have had to carefully review our data collection methods and the ethical dilemmas of data collection within domestic and family environments.

Our study was based around home visits to a broad mix of 10 homes, made up of a total of around 45 people (with occupancy varying slightly over the study). They included families with children, single occupant homes, homes with intergenerational occupancy or guests, and shared homes; participants came from a variety of ethnic backgrounds, occupations, incomes and age ranges, and worked in both traditional nine-to-five working patterns and in shift work, so that different patterns of communication were necessary. These homes are not intended to offer up statistical data, but to capture a varied set of domestic configurations and forms of occupancy that are more or less representative of typical UK domiciles in an unashamedly qualitative study of practice. The initial pre-deployment study was carried out over multiple visits, involving interviews with household members, asking about their communication activities. We asked them to keep a photo-diary of all “messages” made, whether written notes, or objects that were put in a particular place for someone else to see. Participants were also asked to keep a video diary if possible, but all 10 households declined, but did not mind the videoing of their homes at every visit. This research is ongoing, and we have conducted a series of ‘experimental’ technology probes [8] to date.

What has to be recognised is the difficulty in gaining access to homes. Our experience was much different to that described in [9], where the interest for public participation was very high. This possibly has to do with the recruitment method; Crabtree and Rodden had recruited through an advertisement in a national newspaper, whilst in our study, we relied on “snowballing”: recruitment information was circulated via mailing lists to members of staff and students in our department, approximately 500 people, and was also forwarded to other lists by some of them. Although some respondents were interested and got in touch, when discussing it with members of their home they had to pull out. A lot of time was spent in negotiations before finalising the first visit, with one case with a first visit set-up when the family pulled out. Participants who were originally hesitant admitted that they felt uncomfortable with the idea of being asked about how they do things at home and the nature/content of their messaging activities. They were also concerned that some questions would be too personal and intrusive. Usually halfway through the first visit, participants admitted that they had been apprehensive, but relaxed when they realised the actual nature of the study. However, as social relationships built up with the participants, subsequent visits were usually easier.

Another of the concerns that we were particularly worried about was the involvement of minors in the study. Given their heavy use of existing communications technologies, they were likely to be important users of the technology, yet this raised questions as to investigating a) what and to whom they were sending messages, b) how we might question them about messages that were interesting, but perhaps sensitive, especially when their parents might not approve of this ‘illicit’ content, and c) how we might question those children about their activities without putting ourselves in a position of potential accusations about our probity and to reassure potential families that we have responsible motives, whilst at the same time, collecting interesting and useful data about messaging activity. Whilst these are clearly important issues, we have yet to encounter them in the main part (although to an extent c) is still relevant), as we have not evaluated the prototype in homes yet. However, this does have serious implications for us, and it may help to use diary information that is not monitored directly by the parents, although this will necessarily require their consent.

4. DESIGNING NETWORKED IT FOR THE HOME

Here, we cover issues and problems in designing IT for the home environment. These are both interactional (problematic relationships arising between the interface and the particular context of the home), and social (problematic relationships arising from the provision and increased visibility of information within the home). Of course, we recognise that these may be interrelated with one another. We do not necessarily provide solutions here, but suggest where problems lie that may require attention in design.

4.1 Interactional issues

It needs to be remembered that any home-based system is accessible, and may be utilised, by people of different ages, physical and cognitive abilities, and even physical stature. Designing such systems could be enormously problematic, in providing a high enough level of functionality to support complex use, through simple actions at the interface. Here, the information appliance design principle of ‘single in function, open in use’ may be of benefit, allowing simple and non-technical interactions to be built up into patterns and contexts of use that support more sophisticated activities (e.g. [10]). There are other appliance-related issues in networked systems of devices: to give the user a consistent ‘feel’ across the device range, at some level, there should be consistency across input types, and consistency across the media displayed. This has implications for the usability of the system and for the ways that information from the various devices can be combined and integrated together.

In any information-limited system (either storage and/or screen real-estate) that has a shared interface, and for which the content is which is not ‘owned’ by someone, there is the potential for a ‘tragedy of the commons' effect. There are clearly information management issues in the design of any such system, to ensure that a shared screen resource does not become clogged with visual material and overloaded to the point that it is unusable. There are potential ways to resolve these, such as supporting the use of social protocols to ensure responsible action, or to automate an information archival process, but this should be used with caution, as automation may not fit with existing patterns of home life.

Finally, there is an important role for the representation of metadata in contextualising information in the media used. For e.g., when the on_message system receives a SMS text message, the message is visibly 'stamped' with sender information, as a photograph and/or a phone number, the time received, and the type of message that it is (SMS). All of this information allows the readers of that content to interpret it within a context: is it still temporally relevant? Is it of concern to them? Should they treat it as a request for action? Should they reply to the sender?

4.2 Social issues

By making information that would otherwise be held in physical (and consequently, access-limited) repositories, we open this information up to more public access that it would previously have been. Within the home, examples of this includes making a child's private materials visible to their parents, when the child may or may not be fully aware of this. Indeed, it is the nature of much IT technology that this access to content is not fully transparent. As designers, we do need to ensure that where information content is, or will be made publicly available from the networked devices, this should be visible at the point of creation, and users should have the ability to delete content. One way to achieve this (and which is the approach that we have selected) is to open up all content to access by all devices within the home – minimising the risk of this confusion.

Another concern about the social effects of networked IT in the home that concerns us is the potential impact of the technology in facilitating antisocial behaviour, which for e.g. in the family home could involve bullying. These concerns give rise to management issues, not necessarily of content control, but of content management: who can access and remove information held on the system. Of course, this is similar to a paper-based system, many of which we have seen in the homes we visited. However, there is a difference here, in that electronic systems can be more invisible to external monitoring and 'social' policing, and thus are open to what may be deemed as less responsible patterns of use.

The home is not a 'blank canvas' that researchers can place information or technology in without regard for its residents' preferences. There is a real issue here in where to place screens (or other hardware), and homeowners may have strong feelings about where research prototypes or eventual technology might be placed. This may be based on an aesthetic preference, or a pragmatic one in which they do not wish for large objects to be drilled into their walls. In our work, this has led to limits being placed on screen size and display placement, although it remains to be seen whether this is a long term issue: it may be that the perceived value/utility of an eventual system, based on the householders' further experience of the system will lead to accommodation in this respect. Related to this, there is an important point to be made about aesthetics in the home: what is an aesthetic experience (visually, or through the nature of the device interaction) can clash with what we take to be the criteria associated with usability. We can see this in the design of devices such as the mobile telephone and iPod, where the aesthetics of the device form and its interaction methods may override suboptimal interaction designs.

Acknowledgments

This work was funded by Microsoft Research. We would especially like to thank Alex Taylor, Laurel Swan, Abi Sellen and Richard Harper for their help in informing and conducting this research.

REFERENCES

[1] Garfinkel, H. (1967) Studies in ethnomethodology. Englewood Cliffs NJ: Prentice-Hall.

[2] Hindus, D., Mainwaring, S., Leduc, N., Hagstrom, A., and Bayley, O. (2001) Casablanca: Designing social communication devices for the home. Proc. CHI'01, Seattle, WA, USA. p. 325-332.

[3] Vetere, F. *et al.* (2005) Mediating intimacy: designing technologies to support strong-tie relationships. Proc. ACM CHI'05, p. 471-480.

[4] Taylor, A.S. and Swan, L. (2005) Artful systems in the home. Proc. ACM CHI'05, p. 641-650.

[5] Crabtree, A., Hemmings, T., Rodden, T. and Mariani, J. (2003) Informing the development of calendar systems for domestic use. Proc. ECSCW'03, p. 119-138.

[6] Harper, R. Evergeti, V. Hamill, L. and Shatwell, B. (2003) The Social Organisation of Communication in the Home of the 21st Century: An analysis of the future of paper-mail and implications for the design of electronic alternatives. Cognition, Technology and Work, 5, p. 5-22.

[7] Crabtree, A., Hemmings, T. and Rodden, T. (2003) The social construction of displays: coordinate displays and ecologically distributed networks. In: Public and Situated Displays: Social and Interactional aspects of shared display technologies, O'Hara, Perry, Churchill and Russell (Eds). Kluwer, CSCW Series.
[9] Hutchinson, H., *et al.* (2003) Technology probes: inspiring design for and with families. Proc. CHI'03, 17-24.
[10] Crabtree, A. and Rodden, T. (2004) Domestic routines and design for the home. Computer Supported Collaborative Work: the Journal of Collaborative Computing, **13** (2), 191-220
[11] Sharpe, W. P. and Stenton, S. (2002) Information Appliances. In The Human-Computer Interaction Handbook: Fundamentals, evolving technologies and emerging applications (Eds. Jacko and Sears), LEA, Inc.

When Teenagers Tŧype

Janet C. Read, Matthew Horton
Child Computer Interaction Group, University of Central Lancashire
jcread@uclan.ac.uk, mplhorton@uclan.ac.uk

This paper describes what happens when teenagers enter text at a QWERTY keyboard. The study demonstrates where and when input errors were made and classifies them into six types. Most of the errors that were made by the teenagers could to some extent be expected and this raises questions about the spell checking software that is commonly shipped with Word Processing software. The authors propose a modification to spell checking software based on the research findings.

QWERTY, text input, errors, spell checking, teenagers, word processing

1. INTRODUCTION

Despite it being one of the main methods by which users engage with computer technology, the QWERTY keyboard, and the performance of users at it, is not especially well researched. Much of the reported work on text input at the keyboard has used adult performers that have been either trained in keyboarding or were very clearly novices, or has focused on users with disabilities with studies looking at the efficacy of different input methods (Norman and Fisher 1982), (Matias, MacKenzie and Buxton 1996),(Trewin and Pain 1999). Interest in performance at the keyboard waned during the years when the GUI gained influence and although there are several current studies on the efficiency and effectiveness of reduced keyboards, as seen on mobile phones, (MacKenzie and Soukoreff 2002) the lack of research into QWERTY keyboard use is surprising.

One possible reason for the lack of interest in the QWERTY keyboard may be that researchers feel that the QWERTY keyboard, having been a part of technology for so long, is a piece of technology that cannot be changed. However, if only to update some benchmarks and gain some understanding of errors and user behavior, the interaction between users and the QWERTY keyboard still merits investigation.

This paper describes a small study that looked at teenage keyboard behavior. It begins with some background detail that places the study in context and motivates the work. The literature on QWERTY text input is then presented and this is followed by an explanation of the research study that ensued. The results are presented and discussed and the authors offer some ideas for further work as well as some reflections on the study.

1.1 When Teenagers Type

For several years, the leading author has been engaged in research into the use of handwriting for text input. Most of this work has been done with children, but on carrying out studies with teenagers, this author became interested in the particular behaviors of teenagers as users (Read 2005). There are several studies on teenage computer interaction but these mostly focus on either interactions using social communication tools like mobile phones and instant messaging (Taylor and Harper 2002), (Berg, Taylor and Harper 2003), or on design activities (Denham 1993).

With regard to text input, teenagers would be expected to perform better than younger children as their spelling patterns are more settled and their ability to recall short text phrases is improved. Teenagers might also be expected to perform better than some adults in text input tasks as, as schoolchildren, they are typically conscientious about writing and spelling and have a desire to perform well. However, predicting against them is their lack of tutoring in keyboarding for text input and their extensive use of MSN and SMS messaging that has created a special vocabulary.

It is common in research studies of text input to use undergraduate students. These students are in some ways similar to the teenage population considered here, but, their involvement in Higher Education implies a certain level of ability, they are often (by virtue of the text input studies being carried out in computing departments) heavy keyboard users, and they are generally unused to writing with pens or pencils.

In the developed world, the typical teenager (if there is such a species) would write by pen and at keyboards (including reduced key keyboards as found on mobile phones) in roughly equal measure. This teenager would be used to having his or her work judged by its appearance although the metrics used for judgment might vary, for instance in text messaging credit might be given for text speak, whereas in the classroom, only standard native language would be acceptable.

2. INVESTIGATING TEENAGERS TYPING

To begin to find out what happened when teenagers typed, a small sample of eighteen teenagers aged between 13 and 14 from a mixed ability state comprehensive school were selected to carry out some simple text entry tasks. These teenagers came to the research voluntarily, one at a time, and did the work having been brought out of an English class. Half the pupils were from a high ability class, and half from a low ability class. Within these two groups there was no conscious selection of gender or ability level. The research took place in a corridor outside the teaching room; there was no use of video or audio recording, the pupils' names were not gathered and a single researcher observed each instance with all of the data gathering taking place on a single day using a single machine.

2.1 Design

It was decided to use a copying exercise so that text input errors could be easily identified. In line with other similar studies, short text phrases were used that could be easily recalled by the participant. A set of eleven text phrases were chosen for use; nine of these were from the list created by (MacKenzie and Soukoreff 2003), the other two phrases were designed to include almost all the letters in the alphabet. The phrases that were used are summarized in Table 1.

The pupils were placed into three groups, A, B, and C. On attending the experiment, the pupils had the procedure explained to them and were advised that they could stop whenever they wanted and that they did not have to finish the tasks. The phrases to be copied were printed in a size 20 font and presented to the pupils one at a time on rectangular (5cm by15cm) cards. Some pupils placed the cards in front of them, between the keyboard and the screen; others left it on the desk alongside the computer. Most of the pupils appeared to glance at the phrase once, memorise it, and did not refer back to the phrase once they had started to type.

Group	Phrase	Characters
All	The quick fox jumped over the gate	28
All	The lazy brown dog cried	20
A1	Prevailing wind from the east	25
B1	Never too rich and never too thin	27
C1	I can see the rings on Saturn	23
A2	Physics and chemistry are hard	26
B2	Time to go shopping	16
C2	Elephants are afraid of mice	24
A3	The world is a stage	16
B3	Take a coffee break	16
C3	What you see is what you get	22

TABLE 1: The text phrases that were used

Each pupil typed the same first two phrases into the computer and then entered three more phrases depending on which group they were in. Within each group, the order in which they entered these three phrases was shuffled so as to eliminate learning effects from within these three phrases. The phrases were entered using a laptop PC and with Notepad® on the computer. The notepad file was set up so that the text was displayed in a size 14 font.

2.2 Analysis

During the text entry activity, all the keystrokes were logged using an automated logger, this allowed the researcher to capture the mistakes as they were made and was essential as almost all the errors were corrected by the pupils. The time on task was not collected. The pupils were told to take as long as they wanted, and were told to correct as they felt fit.

When all the data had been gathered in, the errors made by the pupils were collated into six categories using an open card sorting method. These categories were, spelling errors (SE), next to errors (NT) close errors (CE), space errors (SC) double characters (DC) and unknown reasons (UR).

Examples of the different errors are given here:
- Spelling error (SE) ‘writing chemisry instead of chemistry’
- Next to errors (NT) ‘ pressing r instead of t’ (next to on the keyboard)
- Close errors (CE) ‘pressing n instead of h’ (defined to be those where the keys were diagonally touching)
- Double characters (DC) ‘writing thinn instead of thin’
- Space errors (SC) ‘writing overt instead of over the’
- Unknown Errors (U) those for which there was no obvious reason

2.3 Results

The pupils in group A were expected to write a total of 115 characters, in Group B, 105 characters, and in group C, 117 characters.

For each pupil, an error rate (ER) was calculated by taking the total number of errors and dividing by the number of expected (optimal) characters. The errors are summarized in Table 2. The mean error rate was 0.0256, the standard deviation was 0.012. This indicates that in general the pupils made only 2 or 3 mistakes per hundred characters.

Examination of where the spelling errors were made revealed that most were made in group B where the phrases ‘never too rich and never too thin’ and ‘take a coffee break’ caused problems as several pupils spelt coffee and too wrong. The biggest single cause of errors was in hitting a ‘next to’ key; half of the pupils making at least one ‘next to’ error. Only one pupil (number 3) managed to enter the five phrases without error.

Group	**SE**	**NT**	**CE**	**DC**	**SC**	**UK**
A	1	7	1		2	4
B	7	4	2	2		2
C	5	7	1			4
Total	**13**	**17**	**3**	**2**	**2**	**10**

TABLE 2: The errors made by the pupils

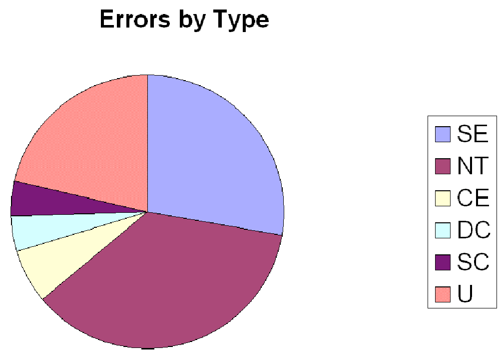


FIGURE 5: Errors by type

The classification of errors into the six types was not without problems. Where a key that was entered was next to the one intended, the assumption was made that this was a next to error; however, this could easily have been a spelling error with it just being a coincidence that the keys hit were next to one another. There was a similar problem with the close errors.

3. DISCUSSION

It is not possible to eliminate errors from text input tasks. Each single error reported in this study cost the typist at least two extra keystrokes (one to delete and another to insert the correct character). Often, there were many more keystrokes as mistakes were seldom noticed on occurrence and correcting them necessitated multiple backspaces and multiple re-keying.

Given that errors will happen, it is interesting to contemplate whether or not is it possible to ‘design’ the effects away, especially given that during typing, users seldom look at the screen after each keypress or even after each word. A common design solution for keyboard errors is spell checking and grammar checking. These tools, shipped with most word processors, are intended to help eliminate and rectify errors, and it is sometimes the case that spellings are automatically corrected

3.1 The influence of the spell checker

Given the relatively common occurrence of 'next to' errors, a further investigation was made into those that occurred in this study. Given that most of the pupils (corrected at the end of the phrase, each next to error resulted in a misspelt word that lingered until found.

Error	Intended word	Word as entered	Spell check list	Position of desired word
1	jumped	junped	Jumped (1)	1
2	prevailing	prevailung	Prevailing (1)	1
3	wind	wimd	Wind (1)	1
4	is	ia	Air, aim,ail,aid, is (5)	5
5	chemistry	chemisrry	Chemistry (1)	1
6	world	Wor;d	Word, world (2)	2
7	**dog**	**sog**	Sag, slog, song, smog, so –	NOT FOUND
8	*time*	*tome*	Nothing offered	NOT SEEN AS A MISTAKE
9	break	nreak	Break (1)	1
10	fox	foz	Fez, Fuzz, Fox (3)	3
11, 12	**mice**	**mivr1**	Mar, mire, mover, miry, mir –	NOT FOUND
13	you	uou	You (1)	1
14	elephants	elephanta	Elephant, elephants (2)	2
15	**you**	**tou**	Too, tour, tout, toe, tofu –	NOT FOUND
16	jumped	jumprd	Jumped (1)	1
17	you	yoi	Yogi, you (2)	2

TABLE 3: The Spell Checker suggestions

These 'words' were entered into Microsoft Word® 2003, and in each case, the offerings from the spell checker were noted. These can be seen in Table 3. The spell check list shows either all the words that were presented (in the order presented) or all the words up to the one that matched the desired word. In auto spelling correction, the first item in the spelling list is used to replace the erroneous word. These results show that for some of the errors ((7, 11, 12, 15) shown in bold), the spell checker did not offer the desired word at all, and it also highlighted ((8) shown in italics) the problem when adjacent keys pressed in error result in a real (albeit incorrect) word. It can be seen in this table that less than half of the mistakes were corrected with the first choice in the spell checker.

This very simple study has shown some of the problems with the traditional spell checker. Spell checking software uses MSD (minimum string distance) algorithms to 'suggest' appropriate words, when a word is incorrect but has the correct leading letter and is of a reasonable length, the software has a high chance of finding a fit but with shorter words, especially those that have an incorrect leading letter, the spell checking algorithms are less effective A particular example is the difficulty found with 'tou' and 'you', there is possibly a case for designing a spell checker that applies a higher weighting to next to typing errors. In this case the words that have an MSD of one away from 'tou', i.e. Too, tour, tout, toe, tofu would include 'you' and this would be placed ahead of 'tofu' as the experimental results here indicate that it is more likely that a next to error (y for t) will be made than a letter be omitted.

4. CONCLUSION

The teenagers in this study used a pick and poke typing method; none used more than four fingers. With a cohort that had had formal keyboarding instruction, the results would have been different. In addition, the spelling mistakes that they made were (according to the class teacher) fairly typical of this age group but would not necessarily occur in groups that had learnt English as a second language and spelling mistakes might not occur with the same density with different language groups.. The choice of text phrases will have affected the results. It would be possible to replicate this study with different phrases and with different results, as, for instance, eliminating 'ambiguous' or 'high risk' words, like 'there', 'coffee', 'too' and 'really' would appear to reduce spelling errors.

Further work is needed to investigate how a differently designed spell checker might help the peck and poke typist, to investigate the effect of different phrase sets on text input experiments, and to determine whether or not this teenage generation of computer user will change their input behaviors as they get older. This latter area of research is especially challenging whilst also being especially interesting. The teenager of today is vastly different from the teenager of ten years ago; the child today will become a different teenager than those in this study. It is interesting to speculate where the QWERTY keyboard will go. The improvement in digital in technologies and the potential for smaller keyboard devices offer conflicting options for text input in the future.

[1] There were two mistakes in this word.

REFERENCES

Berg, S., A. S. Taylor and R. Harper (2003). Mobile Phones for the Next Generation: Device Designs for Teenagers. CHI 2003, Ft Lauderdale, Florida, ACM Press.

Denham, P. (1993). "Nine- to fourteen-year-old children's conception of computers using drawings." Behaviour and Information Technology **12**(6): 346 - 358.

MacKenzie, I. S. and R. W. Soukoreff (2002). "Text Entry for Mobile Computing: Models and Methods, Theory and Practice." Human-Computer Interaction **17**(2): 147 - 198.

MacKenzie, I. S. and R. W. Soukoreff (2003). Phrase Sets for Evaluating Text Entry Techniques. CHI 2003, Ft. Lauderdale, FL, ACM Press.

Matias, E., I. S. MacKenzie and W. Buxton (1996). "One-Handed Touch-Typing on a QWERTY keyboard." Human-Computer Interaction **11**: 1 - 27.

Norman, D. A. and D. Fisher (1982). "Why alphabetic keyboards are not easy to use: Keyboard layout doesn't much matter." Human Factors **24**(5): 509 - 515.

Read, J. C. (2005). The Usability of Digital Ink Technologies for Children and Teenagers. HCI2005, Edinburgh, UK, Springer.

Soukoreff, R. W. and I. S. MacKenzie (2001). Measuring Errors in text entry tasks: An application of the Levenshtein string distance statistic. CHI 2001, New York, ACM Press.

Taylor, A. S. and R. Harper (2002). Age-old Practices in the "New World": A study of gift giving between teenage mobile phone users. CHI 2002, Minneapolis, ACM Press.

Trewin, S. and H. Pain (1999). "Keyboard and mouse errors due to motor disabilities." International Journal of Human-Computer Studies **50**: 109 - 144.

Benchmarking Desirability over Time

Susan Dziadosz and Rive Citron
1 Microsoft Way, Redmond, WA 98052, USA
{sudzi, rivec}@microsoft.com

Product Reaction Cards are a usability technique that is used to elicit users' emotional reaction to products. We describe how this method has been expanded for use outside of the traditional usability lab setting with a case study from the U.S. and Germany. This case study evaluated users' perceptions of handwriting recognition on Tablet PCs over time, in response to periodic software upgrades. We found the method was generally successful at eliciting detailed user feedback during a longitudinal satisfaction benchmarking study carried out simultaneously in both countries. We also point out some of the method's shortcomings when employed in a real world context.

Graphical user interfaces; emotion in HCI, user-centered design; evaluation/methodology.

1. INTRODUCTION

What really sets a product apart from the competition? What are the most salient positive and negative aspects of your product in users' minds? What is your product's emotional impact on users? How can user researchers elicit information about these distinguishing differences and use the information to improve future products?

Questions like these are of increasing interest to the usability community and have been widely explored in the literature. Some authors discuss the role of "emotion" in users' experiences with product design, arguing that both usability and beauty must be considered when designing products [1], [2]. For example, a web site's affective qualities can influence users' perceptions of how useful and usable the site will be [3]. Other authors have concentrated on intangible elements such as "pleasure" [4], "fun," [5], "hedonic quality" [6], and "desirability" [7], [8]. Many researchers are pursuing methods for measuring users' emotional reactions, for example, using thermal imaging as an unobtrusive method for gauging emotional states [9] or methods for eliciting direct verbal feedback from usability participants in a lab setting [7].

This paper will focus on how we adapted one of the tools from the desirability toolkit, namely, the product reaction cards [7]. While previous work centered on collecting data in the usability lab, we adapted the product reaction cards method for use in benchmark research. We will first describe the product reaction card methodology, and then go on to outline the case study that describes how the tool was adapted and the data obtained.

1.1 Product Reaction Cards Method

One method from the desirability toolkit is the product reaction cards. The concept behind the cards is fairly simple. The researcher wants to learn, with as little prompting as possible, what the participants in a usability study really feel about the product they evaluated. But participants may be reluctant to give criticism, or may be unsure of how best to express their opinions. Our experience with using the product reaction cards confirms claims that the cards trigger participants to give detailed and enlightening feedback [7].

After completing the usability evaluation, the participant is given a stack of 118 index cards with words or short phrases on them. They're asked to pick the words that best describe the product or how the product made them feel. After the participant sorts the cards, the researcher returns, records which cards were selected, and then asks the participant to narrow down the set to the five best descriptors. The most critical part of the method is an interactive discussion about why the participants picked the five words or phrases they did, including detailed examples from the product. The cards quickly provide a simple framework for a semi-structured interview with the usability participant. Figure 1 lists some of the words or phrases used in the product reaction cards.

One important thing to clarify is the difference between reporting the words and leveraging the data from the discussion about why users selected a particular word. The word cards are used merely to provide the participants with something to react to; they are not designed to be the only, or even the primary, output from

Accessible	Desirable	Gets in the way	Patronizing	Stressful
Appealing	Easy to use	Hard to use	Personal	Time-consuming
Attractive	Efficient	High quality	Predictable	Time-saving
Busy	Empowering	Inconsistent	Relevant	Too technical
Collaborative	Exciting	Intimidating	Reliable	Trustworthy
Complex	Familiar	Inviting	Rigid	Uncontrollable
Confusing	Flexible	Not valuable	Slow	Unpredictable
Consistent	Frustrating	Overbearing	Stimulating	Useful

FIGURE 1: Selected words from the set of 118 product reaction cards

the tool. The words do not always capture the true feeling of the users toward the product. The most accurate feedback comes from the discussion about why each word was selected. For example, the word optimistic could be interpreted as positive if it's part of a verbatim such as "this is an optimistic view of the future of computing." Or it could mean something very different if the verbatim was "you are very optimistic if you think anyone is going to be able to use this product." We raise this point because it played a role in how we adapted the method for our research, as you will see from the case study.

2. CASE STUDY

Previous field trials conducted by the Tablet PC user research team focused heavily on note-taking behavior and general adoption of the Tablet PC technology [10]. In June 2003 we began another field trial with a year-long focus on the handwriting recognition user experience for twenty U.S. and twenty German knowledge workers who were using Tablet PCs as their main work computers. This field study benchmarked satisfaction in response to periodic software upgrades, such as changes to the handwriting recognizer algorithms, changes to the user interface for entering handwriting for immediate recognition, and a feature that enabled users to adapt the recognizer to their specific handwriting style.

2.1 Handwriting recognition on Tablet PCs

A Tablet PC is a fully functional personal computer that runs Microsoft® Windows® XP Tablet PC Edition. It's equipped with an electromagnetic screen that's designed for interaction with a digital pen. Users use a digital pen, called a tablet pen, to interact with items on the screen, for example, to select, drag, and open them. They can also use a tablet pen instead of a keyboard to handwrite information directly in a program on a Tablet PC, for example, to compose an e-mail message. The Tablet PC can instantly convert the handwriting to typed text, or the user can keep the handwriting as ink.

Handwriting recognition enables users to input text when it's inconvenient or inappropriate to use a keyboard, such as in a class, a meeting, or an interview, or while standing or lying down. Although handwriting recognition technology can make users more productive, inaccurate recognition reduces productivity gains and can be a source of frustration for end users, which may inhibit their use of a Tablet PC. For example, users may avoid using handwriting recognition for tasks that require high accuracy [11].

2.2 Tablet PC multi-country extended field trial

Our U.S. English and German handwriting recognizers use neural networks that are trained on millions of handwriting samples, which we've been collecting for several years from a diverse population. This approach results in most users enjoying good handwriting recognition upon immediate use. The recognizers have been improving continuously over recent years; however, no handwriting recognizer that we're familiar with attains 100% accuracy, so users must still cope with incorrect recognition to varying degrees. But measuring user satisfaction with accuracy is complex because satisfaction is influenced by a number of factors including: the back-end recognition engine performance, the user interface, and factors related to general tasks and context. Hence, we used a triangulation approach that combined three types of methodology to increase confidence in the overall satisfaction measures gained. Two methods were a Likert-scale survey and the NASA Task Load Index [12]; their description is beyond the scope of this paper. The third method was the product reaction cards that were used in one-on-one interviews lasting 45-60 minutes with each participant. When choosing their cards, as shown in Figure 2, participants were always instructed to focus on the period of time since the most recent software upgrade.

The product reaction cards had to be translated into German before benchmarking began in that country. Five bilingual German natives collaborated on the translation effort, including three internal employees who specialized in language-related roles in the company. Some translations were unanimously agreed upon by all five translators, such as *fragile = zerbrechlich*. For other words, they generated several possible alternatives, such as *customizable* = {*anpassbar, benutzerdefinierbar,* or *personalisierbar*}. To reach

FIGURE 2: A field study participant selecting product reaction cards that describe her reactions to the software

consensus, the group met to discuss the nuances of the German suggestions and what was meant by the original English words. Because the goal of the product reaction cards is to assist participants in generating verbose feedback, if a strict translation of an English word did not make sense, the translators used a word that captured the spirit of the English word instead.

2.3 Results

On the whole, the product reaction cards technique proved successful in this German and U.S. longitudinal field study. As recommended [7], when analyzing the interview transcripts, we ignored the words and focused instead on the reasons the participants cited for *choosing* those words. We categorized the feedback according to the product areas driving the positive or negative reactions and were able to generate lists of top “pain points” with the handwriting recognition user interface (for example, the error correction process) and with categories of recognition accuracy problems (for example, numerals or abbreviations). This feedback has been embraced by the feature team that is focused on improving the end-user handwriting recognition experience. We were able to address certain problems in time for the release of Windows XP Tablet PC Edition 2005 and are using the feedback to prioritize user experience improvements for future releases.

However, in addition to the benefits of this technique, we also faced some unique challenges to using the product reaction cards in a real-world context. Some benefits and challenges were specific to the field study setting; others arose from the distributed aspect of the research.

2.3.1 Product Reaction Cards in a Field Study Setting

As expected, the product reaction cards helped the participants in both countries generate informative feedback about the parts of the handwriting recognition experience that they perceived as improving, staying unchanged, or worsening between upgrades. We found no differences between Germany and the U.S. with respect to the ease of using product reaction cards in interviews. The technique was robust enough to allow similar themes to emerge as appropriate in both countries, but flexible enough for emergence of themes that are unique to each country. For example, one similarity was that in both countries, handwriting recognition was viewed as especially useful when the users were mobile. One difference was their reactions to modifying their handwriting to improve recognition accuracy – the U.S. participants were more willing to modify their handwriting than the German participants were.

Each participant reported the experiences that were salient for him or her, which gave the researchers and the product team unparalleled insight into the individual drivers of satisfaction and dissatisfaction with using one’s handwriting to enter text on a computer. We also gained greater understanding of the contexts in which the participants found handwriting recognition more or less valuable in the course of their daily work. Furthermore, because the product reaction cards were used in longitudinal research, we were able to track whether any aspects of the user experience were improving or degrading between upgrades.

However, the handwriting recognition experience was only one facet of the overall experience the participants had with the Tablet PC. The computer hardware was new to everyone, and participants were adjusting to several other new software experiences during the study. As a result, despite being instructed to focus on their handwriting recognition experiences since the last software upgrade, the top five words they chose were not always “on topic.” Particularly in the U.S., most participants occasionally gave feedback about the computer hardware, the note-taking software they were using, and how mobile computing affected their interactions with colleagues and clients—just to name a few.

This additional information seemed to be a direct consequence of the strength of the product reaction cards as a “Top of Mind” technique. At first, this strength might seem a double-edged sword – perhaps it is not effective for evaluating one component, such as handwriting recognition, of an overall novel system, such as a Tablet PC. Our opinion is quite the opposite, however. When feedback was not specifically about handwriting recognition, we realized this could signify a few things. One possibility was that at a high level, other experiences were more important or valuable to the participant than the handwriting recognition features. This information could prove crucial to understanding the overall value proposition of the whole system and how customers think the various experiences fit together.

A second possibility was that the latest upgrade had not made enough of a difference to that particular participant's experiences. To explore this possibility, the interviewer asked if the participant had additional feedback about the upgrade. For example, the interviewer asked a general question such as, "Has handwriting recognition gotten worse, stayed the same, or improved for you?" Or the interviewer referred to previous interview notes to ask a specific question such as, "Last time you mentioned that you were having difficulties getting the letter *k* to be recognized accurately. Has anything changed with respect to the letter *k* since the last upgrade?"

2.3.2 Product Reaction Cards in Distributed Research

Because this field study was being carried out simultaneously in two locations by two teams of researchers, there were some differences in how the interviews were conducted and the subsequent data was collected. Because a laptop is sometimes perceived to be a social barrier, in the U.S. the researcher asked the participants' permission to type notes during the interviews. Hence, this afforded the researcher the opportunity to record a large volume of data in great detail.

However, prior research suggested that using a laptop during a conversation is considered even more socially unacceptable in Germany than in the U.S. Consequently, the German researcher was required to take handwritten notes of what he believed were the most important points participants made about each of the chosen top five cards. Consequently, different volumes of information were generated in each country. A typical U.S. interview transcript might contain more than 5,000 words compared to 500 words for a German interview transcript from this project. So, although the themes that emerged from the interviews in both countries were very similar, there was a greater amount of illustrative examples in the U.S. transcripts, which provided more information about what each chosen word typified for the participants.

3 CONCLUSION

We have described a case study that used the product reaction cards outside of the traditional usability lab setting and internationally. We found the method was successful in longitudinal satisfaction benchmarking conducted in the U.S. and Germany. It provided the same benefits that have been reported in U.S.-based usability lab studies, namely that users are encouraged to provide candid feedback rapidly, users do not hesitate to give negative feedback, and the product team finds the users' feedback useful for improving the product.

Researchers are cautioned to take care in overseeing the translation of the product reaction card words into another language, because cultural nuances can strongly affect the interpretation of a given word. Additionally, when using the cards to elicit feedback about a component of a larger system, researchers should be prepared for user comments to stray from the desired topic. Interviewers must be adept at accepting off-topic feedback while finding ways to steer the users back to the topic under evaluation. Finally, researchers should consider the social acceptability of the various methods of note-taking during interviews for a given locale. For example, in cases where it is deemed unacceptable to type notes, interviewers should consider capturing data using audio recordings, which can be transcribed at a later date.

While we continue to refine these techniques, we hope other researchers will incorporate techniques to measure desirability in their evaluations and to share their results.

ACKNOWLEDGMENTS

The authors wish to thank Joey Benedek and Trish Miner for many discussions about the product reaction cards.

REFERENCES

[1] Norman, D. A. (2004) Emotional Design: Why We Love (Or Hate) Everyday Things. Basic Books, New York.

[2] Desmet, P.M.A. & Hekkert, P. (2002) The Basis of Product Emotions. In Green, W. and Jordan. P. (eds.), Pleasure with products, beyond usability. Taylor & Francis, London.

[3] Zhang, P. & Li, N. (2005) The Importance of Affective Quality, *Communications of the ACM*, **48**(9), 105-108.

[4] Jordan, P.W. (2000) Designing Pleasurable Products. Taylor & Francis. London.
[5] Blythe, M. A., Overbeeke, K., Monk, A. F., and Wright, P. C. eds. (2003) Funology: From Usability to Enjoyment. Kluwer Academic Publishers, Netherlands.
[6] Hassenzahl, M., Platz, A., Burmester, M., & Lehner, K. (2000) Hedonic and Ergonomic Quality Aspects Determine a Software's Appeal., Proceedings of CHI 2000 Conference, The Hague, Netherlands, 01-05 April, pp.201-208. ACM Press, New York, NY, USA.
[7] Benedek, J. and Miner, T. (2002), Measuring Desirability: New Methods for Evaluating Desirability in a Usability Lab Setting. Proceedings of Usability Professionals Association 2002, Orlando, FL, 08-12 July.
[8] Saunders, E. B. N. (1993) Product Development Research for the 1990's. Proceedings of the Human Factors and Ergonomics Society 27th Annual Meeting, Seattle, WA, 11-15 October, pp.422-426.
[9] Puri, C., Olson, L., Pavlidis, I., Levine, J., and Starren, J. (2005) StressCam : Non-Contact Measurement of Users' Emotional States Through Thermal Imaging. Proceedings of SIGCHI 2005, Portland, OR, 02-07 April, pp. 1725-1728. ACM Press, New York, NY, USA.
[10] Feldman, E., Pennington, E., and Ireland, J. (2003) Tablet PC – Using Field Trials to Define Product Design. Proceedings of HCI International 2003, Crete, Grece, 22-27 June, pp.636-640. Lawrence Erlbaum Associates, Mahwah, NJ, USA.
[11] Frankish, C., Hull, R., and Morgan, P. (1995) Recognition Accuracy and User Acceptance of Pen Interfaces. Proceedings ACM CHI 1995, Denver, CO, 07-11 May, pp.503-510. ACM Press, New York, NY, USA.
[12] Hart, S.G., & Staveland, L.E. (1988) Development of a NASA-TLX (Task load index): Results of empirical and theoretical research. In P.A. Hancock and N. Meshkati (Eds.), Human Mental Workload, 139-183. Amsterdam: North-Holland.

Sketch Radar: a Novel Technique for Multi-Device Interaction

Dzmitry Aliakseyeu and Jean-Bernard Martens
Department of Industrial Design, Eindhoven University of Technology
P.O.Box 513, 5600 MB, Eindhoven, The Netherlands
{D.Aliakseyeu;J.B.O.S.Martens}@tue.nl

A plethora of reaching techniques, intended for moving objects between locations distant to the user, have recently been proposed and tested. One of the most promising techniques is the Radar View. Up till now, the focus has been mostly on how a user can interact efficiently with a given radar map, not on how these maps are created and maintained. In this paper, we present Sketch Radar, a system created with off-the-shelf components that allows the user to create and modify radar maps in a flexible way.

Interaction techniques, reaching, large-display systems, multi-display systems, barcode reader.

1. INTRODUCTION

In response to the rapidly reducing cost of display and network technologies, situations in which devices with heterogeneous display sizes interact together are becoming commonplace. Often these environments present a mixture of personal devices such as PDAs and shared devices such as large displays. In a device-cluttered space the tasks of identifying a particular device and facilitating the transfer of objects from one device to another, also referred to as multi-device (display) reaching, become frequent. Alternative techniques for performing such interactions have lately received a fair share of attention in research.

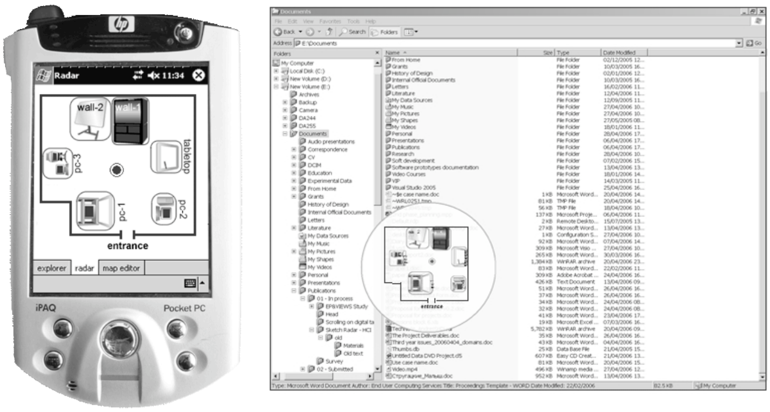


FIGURE 1: Sketch Radar on a PDA and within file explorer on a TabletPC or desktop computer. The layout of the map mirrors the physical locations of devices in a laboratory. The red dot in the middle represents the user or avatar.

In this paper, we propose and describe Sketch Radar (see Figure 1). It is a novel interaction technique that is based on the radar metaphor. With it, a user is able to control how and what information is presented on the radar at any time. The representation of a device on the radar map can be acquired in a direct and explicit way. In the current prototype this is accomplished by means of a barcode reader that reads a barcode of the device. Subsequent interactions can occur through the representation of the device on the map.

2. RELATED WORK

The related work can be subdivided into several parts: 1) multi-display reaching and interaction techniques for large displays, 2) interaction techniques for connecting and identifying devices, 3) remote control techniques.

Several interaction techniques for intuitive and efficient reaching between devices have recently been proposed. In addition, a number of techniques for large wall and tabletop interaction, such as Drag-and-Pop [1] and Push-and-Pop [4], can also be adopted for the purpose. In a recent study, Nacenta et al. [8] found that Radar View, a technique in which the user can pinpoint the desired source and destination on a reduced map, performed significantly better than related techniques like the Pantograph [4,8] and Pick-and-Drop [10]. The spatial overview over displays (devices) in Radar View allows for efficient access. It was also shown that the spatial organization of displays (devices) allows efficient access to them, in the sense that it outperforms existing tree- or list-based approaches (such as File explorer or Favorites in Internet explorer) [5].

The success of map-based techniques such as Radar View [8] relies on being able to associate a physical device with its representation on the map. However, how this association is accomplished and maintained has, as far as we know, rarely been discussed. Usually this process is hidden behind a “smart system” that knows at all times what to present on the map, including how and where all objects should appear. The only known example of a system that uses the radar metaphor and that addresses how physical devices can be arranged on a map is ARIS [2, 3]. ARIS uses an iconic map of a space as part of an interface for performing application relocation and input redirection tasks. The differences with the proposed Sketch Radar are the following. First, Sketch Radar aims at supporting different tasks, i.e. placing and retrieving files (i.e., a reaching task). Second, Sketch Radar is not limited to devices that have screens, but can include devices such as printers. Third, because Sketch Radar does not necessarily rely on a physical layout, such as the devices in a single room, it allows combining distant devices in a single map. Fourth, the nature of the tasks and spaces in ARIS implies that the flexibility in map layout offered by Sketch Radar is not required.

A second class of interaction techniques, such as SyncTap [11], Proximal Interactions [12] and InfoPoint [6], aim at identifying devices in a direct and explicit way. A shared disadvantage of these techniques is that users need to be physically close to the devices in order to perform identification and interaction. This will be advantageous, but not essential, in Sketch Radar.

A third class of interaction techniques, i.e., Semantic snarfing [7] and the Personal Universal Controller [9], aim at using a PDA or other mobile device as mediator between stationary devices. The PDA becomes a (remote) control, especially for those devices that do not possess external controls or a display. Most of these techniques have similar limitations to the techniques for device identification. The interactions are direct and explicit, but require the user to be in front or close to the devices, which usually implies that interaction with only one device at a time is supported.

3. SKETCH RADAR

The use of Radar Views is not limited to “exchanging” objects between distant devices. The Radar technique can also be used for establishing a connection with a specific device. This device may be physically out of reach and need not have its own external controls (for example, a wall display). Establishing such a connection can of course also be accomplished with more direct techniques such as pointing. However, such physical pointing often requires a more complex (and custom-made) hardware setup. Radar Views, on the other hand, can be used in almost any situation and do not rely on a direct line of sight to operate devices.

The common implementation of Radar Views is based on the physical positions of interacting devices. This raises several questions/issues: 1) how is the relevant information (device name, position, size) needed to construct the radar map acquired from the interacting devices; 2) which devices should be presented on the map (should all devices be equally prominent); 3) how should devices be represented (for instance, how can horizontal and vertical screens be represented on a single planar map); 4) what are the boundaries of the map; 5) how to deal with the fact that the map needs to be presented on a screen with limited size and resolution? With the concept of the Sketch Radar we try to address the majority of these questions.

The Sketch Radar consists of three parts: 1) the information gatherer, implemented using a barcode reader, 2) the map editor used for modifying radar maps, and 3) the actual radar for selecting and (re)positioning objects and remote control of devices.

In order to add a new device to the radar map, a user can either “scan“ the barcode on the device, or enter its unique identification name. When a new device is entered it is positioned in the center of the map. By default every device is represented by an image and a text label. The user can reposition and resize the image and label on the current map. The action of entering a device hence enforces the mental link between the real device and its representation on the map. The user can add additional elements, such as text, drawings and images, to this map. These elements are passive, in the sense that they have no influence on the radar interaction, but may be used to clarify the function and layout of a specific radar map.

In a new environment it is usually wise to start with a map that is based on the physical position and size of the devices. After some time, an environment becomes more familiar and tasks become clearer. This may lead the user to readjust the positions, sizes and representations of the devices that are in the map. For example, frequently-used devices may be increased in size and placed closer to the center of the map. Also, by allowing the user to add “sketches” (lines, text) to the map, he can add elements that further strengthen the association between a specific map and a particular task. This flexibility makes the Sketch Radar useful in a range of situations, ranging from interaction in an unfamiliar space, where a close correspondence with the physical arrangement is needed to identify individual devices, to frequent and long-term usage, where the physical space is well known and users can profit from a map that is specifically tailored to their purpose. It is expected that this diminishing importance of "physical" correspondence will go hand-in-hand with a growing user knowledge about the task and space.

Once the radar map has been populated with devices (Figure 2, left), the actual radar interaction itself can be activated. The radar supports two modes of interaction, i.e., discrete and continuous positioning.

In this case the radar interaction is initiated by tap-and-hold (or right mouse click) from the local file explorer application (on the PDA, tabletPC or desktop computer), then a file has already been selected. In the discrete interaction mode, the selected file can be transferred to any of the devices that are represented on the map by simply pointing at the device with the pen. In the continuous interaction mode, the user drags the red dot (which can be seen as an avatar of the user) to the receiving device that s/he wants to interact with. The screen (desktop) of the receiving device appears and the radar application takes control over the mouse of the receiving system (Figure 2, right). In this way, the selected file can be dropped anywhere on the receiving system. When being in the physical neighborhood of the receiving device, the user can also observe the mouse actions on the screen of this system (i.e., functionality similar but simplified to that in Semantic snarfing [7]). This continuous mode can also be used to operate a graphical user interface to a device, even in cases where such a device does not possess a graphical user interface of its own. In such cases, the radar functions as a remote control to the device.

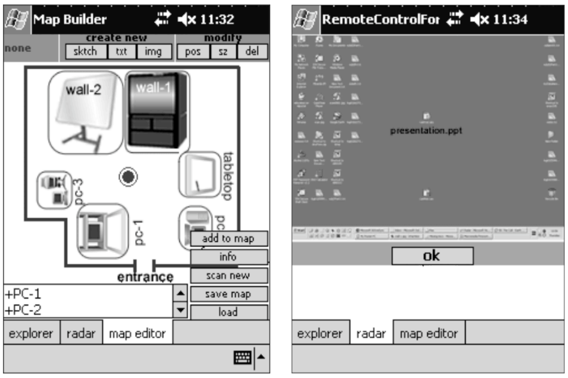


FIGURE 2: Sketch Radar on the PDA: map builder interface on the left, and remote control interface (radar) on the right.

The radar interaction can also be activated directly, i.e., not from within file explorer. The continuous interaction mode can then be used to select one or more files on a remote device. Upon closing the continuous access to this source device, one re-enters the radar map. Either discrete or continuous interaction with a destination device can then be used to transfer, and possibly position, the selected file(s). The PDA thus functions as a mediator between two networked devices.

4. EXPLORATORY EVALUATION

The usability and usefulness of the Sketch Radar prototype was tested in a short exploratory user study. One of the main questions that we wanted to answer was: will subjects change the layout of the map in response to the tasks they are asked to perform? The experiment was conducted with 5 subjects (2 females and 3 males) between the ages of 23 and 49. All subjects had previous experience with graphical user interfaces. They were tested individually.

The experiment comprised of three parts. In the first part subjects received a PDA with a preloaded map of the laboratory room (Fig. 2) in which the test was conducted. All systems represented on the map had about equal size and their position corresponded roughly to their actual position in the room. The subject's task consisted of moving several files from one computer to another while being inside the room. At the start of the second part of the experiment, subjects were allowed (but not forced) to modify the original map, taking into account that they were going to be requested to repeat the interactions in the first part. They were also allowed to continue changing the map during the remainder of the experiment. In the last part of the experiment, two new devices needed to be introduced – a desktop computer and a printer. Before performing this third part, subjects were instructed to scan and put these new devices into the map.

The evaluation showed that users did indeed change the layout of the map to make it more suitable for the particular task that they needed to perform. The map was not only changed at the start of the second part, when subjects were instructed about this feature, but also during the rest of the experiment. Most subjects did not remove devices that were not used, but simply made them small. The most-frequently used devices, on the other hand, were made bigger. The resulting maps hence still resembled the original physical organization, but increasingly distorted. Most subjects added text labels to the radar maps in order to assist them with the tasks. Because the images on the PDA were not clear enough for reliable identification, text labels were considered important by most subjects. This problem might be less prominent in prolonged use, and might also be reduced by using iconical representations (see figure 2) instead of photographs of the systems.

Subjects made very few errors. The errors were not connected with identification, but reflected confusion between discrete and continuous (desktop) positioning. For example, when the task was to position the file into the default location on the computer (discrete), which can be accomplished by just clicking the representation of the target computer, users instead dragged their avatar to that computer and positioned the file (continuously), and vice versa. Only one subject commented that the idea of “placing or dragging myself” on a specific device, to get access to it, was not clear. Another subject commented that it is “easy to think about objects [devices] in a physical way”.

The results of a usability questionnaire indicated that the Radar was easy to use, flexible and understandable (all subjects gave positive ratings). All respondents except one (who had a neutral judgment) agreed that the Sketch Radar allowed them to transfer files quickly.

5. CONCLUSIONS AND FUTURE WORK

The Sketch Radar was shown to be an effective and flexible tool for both performing multi-display reaching, and for constructing and maintaining the radar maps needed for this interaction. The acquisition of the device representations by means of the barcode reader turned out to be very reliable and quick, and most subjects judged it as being fun.

The first prototype was implemented on a PDA and clearly reflected some of the obvious PDA restrictions. Due to the limited size and quality of the PDA display, the images that represented the devices could not be clearly recognized, as well as elements of the remote desktop during continuous interaction. As remarked earlier, it may therefore be advisable to replace the photographs by pictorial representations that are easier to recognize. The problem with remote desktop can be approached in a way similar to Semantic snarfing [7], where all interface parts of the remote desktop where adopted to make it easer to control them from PDA. Another possible adjustment to the prototype is to use a radar map that is bigger than the PDA screen, which can be panned (and zoomed). This would however complicate the interaction, and therefore should be considered carefully.

Our concept and prototype obviously did not address the privacy issues that are also involved in multi-device operations. In case non-accessible devices show up in the radar maps, the most straightforward response would be to simply remove or minimize them. Using a different representation for systems that are only accessible for reading might also be an option.

Another aspect that is clearly relevant within the multi-device environments that we consider here is multi-user collaboration, either co-located or not. Although it is allowed to have multiple Sketch Radar devices operating within a single environment, where subjects can even exchange radar maps, it is less clear how conflicts should be handled and how performance and appreciation should be measured.

The Sketch Radar solves a key problem of the existing Radar interaction technique by providing an easy and quick way to manage one or more maps of available devices. More testing with a larger number of users, who make a more prolonged use of the system, will also be required before the full value of the system can be judged adequately.

REFERENCES

[1] Baudisch, P., Cutrell, E., Robbins, D., Czerwinski, M., Tandler, P., Bederson, B., and Zierlinger, A. (2003) Drag-and-pop and Drag-and-pick: Techniques for accessing remote screen content on touch- and pen operated systems. In Proc. of Interact 2003, IOS press (2003), 57-64.

[2] Biehl, J.T. and Bailey, B.P. (2004) ARIS: An Interface for Application Relocation in an Interactive Space. In Proc. of Graphics Interface, 2004, 107-116.

[3] Biehl, J.T. and Bailey, B.P. (2005) A Toolset for Constructing and Supporting Iconic Interfaces for Interactive Workspaces. Interact, 699-712.

[4] Collomb, M., Hascoët, M., Baudisch, P., and Lee, B. (2005) Improving drag-and-drop on wall-size displays. In Proc. of GI 2005, (2005), 25-32.
[5] Czerwinski, M., van Dantzich, M., Robertson, G.G. & Hoffman, H. (1999). The contribution of thumbnail image, mouse-over text and spatial location memory to web page retrieval in 3D. In Proc. of Interact '99, IOS press, 163-170.
[6] Kohtake, N., Rekimoto, J., Anzai, Y. (2001) InfoPoint: A Device that Provides a Uniform User Interface to Allow Appliances to Work Together over a Network. Personal and Ubiquitous Computing 5 , Springer, 264-274.
[7] Myers, B., Peck, C.H., Nichols, J., Kong, D., Miller, R. (2001) Interacting At a Distance Using Semantic Snarfing, In Proc. of UbiComp'2001, ACM Press, 305-314.
[8] Nacenta, M.A., Aliakseyeu, D., Subramanian, S., and Gutwin, C. (2005) A comparison of techniques for Multi-Display Reaching. In Proc. of CHI 2005, ACM Press, 371-380.
[9] Nichols, J., Myers, B. (2003) Studying the Use of Handhelds to Control Smart Appliances. In Proc. of ICDCS'03, 274-279.
[10] Rekimoto, J. (1997) Pick-and-Drop A Direct Manipulation Technique for Multiple Computer Environments. In Proc. of UIST 1997, ACM Press, 31-39.
[11] Rekimoto, J., Ayatsuka, Y., Kohno, M. (2003) SyncTap: An Interaction Technique for Mobile Networking. In Proc. of Mobile HCI 2003, September 2003.
[12] Rekimoto, J., Ayatsuka, Y., Kohno, M., Oba, H. (2003) Proximal Interactions: A Direct Manipulation Technique for Wireless Networking. Proc. of Interact'03, IOS press.

Simulated Lifting with Visual Feedback

Faieza Abdul Aziz[1] and Ieuan A. Nicholas[2]
[1]Department of Mechanical and Manufacturing Engineering, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia
[2]Welsh e-Science Centre, School of Computer Science, Cardiff University, CF24 3AA, UK
[1]faieza@eng.upm.edu.my [2] i.a.nicholas@cs.cardiff.ac.uk

Improper manipulation of real-world objects increases the risk of developing work-related back injuries. In an effort to reduce such risk and encourage appropriate lifting and moving methods a Virtual Environment (VE) was employed. The VE made use of visual feedback techniques to allow a person to estimate the forces acting on their lower back whilst moving an object from one location to another. This work investigates various types of visual feedback techniques to support lifting operation in VEs. The results reveal that the combined visual feedback techniques performed better than single feedback techniques. This paper also examined and evaluated a "weight perception test" to observe whether the users can identify the weight of a box using the feedback provided in real-time. Combi feedback was used as it was found to be the best method of visual display feedback in warning the users of their ergonomic lifting condition. Users were able to perceive and react to the changes in feedback as the virtual weights were exchanged.

Lower back pain, manual lifting, virtual environments, visual feedback

1. INTRODUCTION

Major causes of injuries to the lower back are the manual materials handling (MMH) tasks of lifting, lowering, pushing, pulling, holding and carrying. Some aspects of these tasks can be designed using the NIOSH Lifting Equation [1, 2]. Research has been conducted on lifting techniques [3, 4], but the use of virtual environments (VEs) to simulate such tasks would give several important benefits. These include ergonomic improvements in the design of factory layouts, manufacturing cells, production equipment, and the implementation of new methods for educating and training employees, such as production operators. Ergonomic lifting is crucial to avoid lower back pain (LBP) and other injuries. In this paper we are concerned with providing real-time visual feedback to indicate lower back stress experienced by the user during a manual lifting task. Real-time feedback allows a user to make adjustments whilst a task is in progress. This paper also investigated a "weight perception test" to observe whether the users can identify the weight of a box using different types of feedback.

2. REAL-TIME VISUAL FEEDBACK

Several factors need to be considered if visual feedback is to be employed to provide real-time information in a virtual lifting task. The first factor involves the degree to which feedback is integrated within the environment. Full integration occurs when the feedback takes the form of a 3D object that is embedded within the environment; one example of this is an object that shows the prescribed position during a virtual lifting task. In the introduction to the Applications Manual for the 1991 Revised NIOSH Lifting Equation [1] it says that lower back pain and injuries attributed to manual lifting activities are among the leading occupational health and safety issues facing preventative medicine. Based on available statistics, almost half of all lower back injuries are related to lifting, about another 10 percent are associated with pushing and pulling activities and another 6 percent occur while holding, wielding, throwing, or carrying materials [5].

In order to maintain a healthy back and to prevent work-related back pain and back injury, ergonomic principals have to be adopted. Ergonomics is a science concerned with the 'fit' between people and their work, and is typically known for its application in solving physical problems at work [6]. It is also used to evaluate the capabilities of the body in relation to work demands. Ergonomic analysis should allow the user to employ 3D and virtual reality (VR) simulations to determine the comfort and safety of factory and office workstations through the design of better workplaces and developing optimized product development cycles. The NIOSH Lifting Equation would be used as guidance to measure the recommended weight limit (RWL), the maximum weight that should be lifted for a particular task and the lifting index (LI). LI expresses the weight that will actually be lifted (the load weight) as a ratio of the RWL. The equations are as follows (see Eq. 1 and 2).

RWL = LC x HM x VM x DM x AM x FM x CM (Eq. 1)
LI = L / RWL (Eq. 2)
where:

RWL = Recommended Weight Limit
LI = Lifting Index
L = Load weight
LC = Load Constant (23 Kg; the weight that all workers are assumed to be able to lift under optimal conditions)
HM = Horizontal Multiplier (calculated from distance in front of worker)
VM = Vertical Multiplier (calculated from height of origin or destination of lift)
AM = Asymmetry Multiplier (1.0 for lifting in the sagittal plane)
FM = Frequency Multiplier (calculated from lifting rate)
CM = Coupling Multiplier (e.g., type of handles on object being lifted)
DM = Distance Multiplier
D = Vertical distance between the origin and destination of the lift (cm)

2.1 Experimental Set-up and Procedure

The VE software is a C-based application that was designed and programmed by the author using CAVELib API. An Onyx 300 visualization server was used to generate the images on a Portico Workwall (a large-scale display device). Stereoscopic 3D images were created through the use of LCD shutter glasses with a refresh rate of 120Hz (60Hz update for each eye). Tracking for head and box position and orientation was performed using six degrees of freedom sensors together with Trackd software. Detail of the system architecture is given in Figure 1. The experimental design had four experimental conditions; one was experiment with no feedback and the remaining three conditions used “Visual” feedback techniques which displayed in real-time. The three types of visual feedback which were used in this study were: Text, Colour and Combi (combination of colour and text). There were twenty subjects who participated in this study, none of them had had any previous experience of VR.

For trials with No feedback, the user does not see any feedback relating to their LI value. In trials with Text feedback, the user receives feedback in Text on their LI results in the format of the lifting grade designated as “SAFE”, “DANGER” or “HARMFUL”. For trials with Colour feedback, the box would change colour according to the LI values. Three colours were chosen: Green representing a Safe Lift, Yellow representing a Dangerous Lift and Red representing Harmful Lifts. In Combi feedback trials the user is exposed to both Colour and Text feedback simultaneously. Users were asked to perform ten trials for each of the four conditions of feedback. A trial consisted of lifting a box from a lower shelf (Shelf 1) to an upper shelf (Shelf 2). The box dimension was 30cm wide, 15cm deep and 40cm long and fitted with handles. It can be assumed for all experiments described in this paper that the virtual box was intentionally made large enough for the user to see the visual feedback irrespective of lifting location. Figure 2 shows one participant performing the lifting task in a VE.

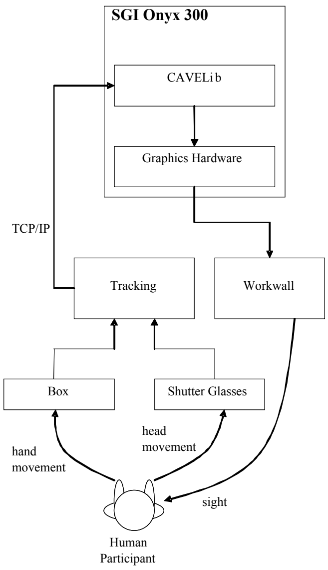


FIGURE 1: System Architecture

FIGURE 2: Participant conducting a trial for Colour Feedback

2.2 Results
Results from the trials were extracted and processed. A one-factor (technique) ANOVA was used for analysis of Task Completion Time (TCT), Percentage of Harmful Lifts (PHL) and Response Time to Feedback (RTF).

2.2.1 Task Completion Time (TCT)
The time taken to successfully accomplish each task was measured. The trials conducted without feedback showed the shortest Task Completion Time. This is probably due to the fact that the participants did not have to monitor any feedback regarding the forces acting on their lower back. This may result in potentially harmful lifts if an improper lift is performed. The results showed that the three feedback conditions all had a similar impact on the Task Completion Time [$F(3,76)=2.35$, $p=0.01$]. However, Combination colour and text feedback gave better results in comparison to both colour and text feedback techniques (mean = 8.7. s.d. = 6.6).

2.2.2 Percentage of Harmful Lifts (PHL)
Results from ANOVA analysis showed that feedback technique had a major effect on PHL. A post-hoc Tukey test reveals that the percentage differed significantly between Text and Combi ($P < 0.05$), and between No feedback and all of the techniques. The plotted graph also shows that Combined colour and text gave the lowest PHL with mean = 3.45 and s.d. = 0.83. Colour feedback outperformed Text feedback with mean = 4.35 and s.d. = 1.09 and mean = 5.55 and s.d. = 1.2, respectively.

2.2.3 Response Time to Feedback (RTF)
Even though the ANOVA results for Response Time to Feedback (RTF) did not reveal any significant difference between the experimental conditions [$F(2,57) = 1.47$, $p < 0.05$], the average for Combined colour and text was the best in performance (mean = 0.46, s.d.= 0.38) compared to Colour feedback (mean = 0.5, s.d.= 0.45) and Text feedback (mean = 0.66, s.d.= 0.38).

2.3 Discussion
Combi feedback displayed a consistently good result. The analysis indicating that this was the best feedback technique for the aspects, TCT, PHL and RTF. The ranking for the results also followed the same path where Colour feedback was better than Text Feedback. Despite the fact that no significant difference was found in TCT and RTF, Combi seems to be the most effective method for alerting the user of their lower back condition while carrying out a manual lifting task.

Even though the No feedback condition took the shortest time to complete the task in TCT analysis; it will result in a greater percentage of harmful lifts. This is a potentially dangerous situation for people, as the symptoms of lower back pain are not normally discovered during the task, but at sometime in the future. So trying to avoid poor lifting technique in the first place is crucial.

3. WEIGHT PERCEPTION TEST
This experiment was designed to measure the participating user's understanding of virtual display feedback. Five trials were prepared, in which boxes were allocated with one of three virtual weights, these were 5kg, 8kg and 12kg and named light, moderate and heavy, respectively. The order in which these weights were used was randomised and the users were not told beforehand which weights had been used in their particular trial.

3.1 Experimental set-up and procedure
The experimental hardware was similar to the previous experiment. The software was another CAVELib application which was programmed by the author. The same sensors were used together with the same box. Only the virtual weight applied to the virtual box was varied. Five trials were conducted with three different virtual weights. Only one type of visual display feedback was employed for this experiment and this was Combi feedback. Users carried out the task individually. The user was asked to lift the box from the lower shelf and place it on to the upper shelf. They were told that they would have to guess the weight of the box according to the virtual feedback received. They were required to give the answer verbally as soon as they noticed the feedback difference or at the latest within 30 seconds of completion of the lifting task. The experimenter first demonstrated how the experiment worked and users were shown the feedback for the three different weights. Once the users had completed the experiment, they were required to fill in a questionnaire regarding the weight perception test which had been done. Users rated the extent to which they perceived each feedback using a 7-point scale ranging from 1 (not at all) to 7 (completely). Higher scores indicate greater perception of weight differentiation.

3.2 Results

The principle basis for this experiment was to evaluate the user's understanding and reaction to visual feedback given during a lifting task, according to the different weight attached to it. The percentage of correct-incorrect selection was calculated and analysed. From the results collected, it was found that 96% of the answers given by the users were correct while only 4% were incorrect. The incorrect answers came from two different users, where the first user gave 3 incorrect answers and the other only gave one. It has been explained by the user who gave three incorrect answers that he did not pay full attention during the demonstration. Therefore, this might be considered as an anomalous case. Results from the questionnaire were analysed and it showed that 90% of users chose between score 5 to 7, where 7 represents the highest (the most noticeable difference of feedback between light, moderate and heavy weight). Only two users chose score 3 and 4 respectively. None of the users choose rate score 1 or 2. All of the users gave their response before completion of the lifting task. Details of the percentage is depicted in Figure 3.

3.3 Discussion

Users were able to perceive and react to the changes in feedback given when the experimenter exchanged the virtual weight during the trials. . It was interesting to note that all the results were given during the lifting task. The Combi feedback which was used for the weight perception test was clear and easy to follow. This was proven by the fact that the users dealt with the feedback changes even without practising the virtual weight changes beforehand.

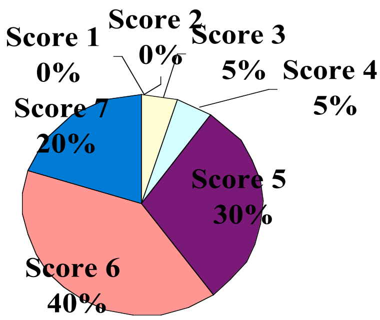


FIGURE 3: Percentage of Weight Perception Test

4. GENERAL DISCUSSION

The overall purpose of this study was to investigate whether or not visual display feedback can be used as an aid to help users monitor their lower back pain whilst performing a manual lifting task. Based on the results of this study, it is recommended that with a good selection of visual feedback, the user might perform well, according to the LI values calculated by NIOSH, given through information in text messages. Combi was the best in performance compared to Colour and Text feedback. Users found that Combi was easy and helpful because they could control for coarse and fine LI. For example, if the user needed to bring the LI value much lower (coarse control), he/she might rely on Colour changes. However if only small changes (fine control) were required, the user would prefer to use Text as this was much more accurate.

The study also verified that users' perception of the differing feedback in the weight perception test was very high. This suggests that users found it easy to understand the feedback even when various weights were applied which changed the LI as well as visual display feedback. The majority would prefer Combi as the best visual feedback followed by Colour feedback and Text feedback. Nonetheless, Text feedback may still be a useful and required feedback cue in a VE design or visualization application where details of numeric numbers are the main goals (and Task Completion Time is not the main objective).

5. SUMMARY

Visual display feedback has been proved to aid users in carrying out manual lifting tasks safely. In order to monitor the forces acting on a user's lower back while performing manual lifting task, the NIOSH equation, which calculates RWL and LI, was applied as a guideline to categorise the lifting regions. The visual feedback displays the changes according to LI values calculated in real-time from sensor data. Three types of visual display feedback have been tested these were Colour, Text and Combi. All of the feedback conditions were suitable for application in a manual lifting task, but Combi was found to be the best according to the results of TCT, PHL and RTF. Combi feedback was also tested on users' weight perception where the majority (96%) gave correct results. Further research is required to determine whether other mixed visual feedback methods actually improve user performance in visualization VEs.

REFERENCES

[1] Waters, T. R., Putz-Anderson, V. and Garg, A. (1994) Applications Manual For the Revised NIOSH Lifting Equation, Atlanta, U.S.A.: Springfield, VA.

[2] Anderson, C. K. and Chaffin, D. B. (1986) A biomechanical evaluation of five lifting techniques. Applied Ergonomics, **17**, p. 2– 8.

[3] Bobick, T. G., Belard, J. L., Hsiao, H. and Wassell, J. T. (2001) Physiological Effects of Back Belt Wearing During Asymmetric Lifting, Applied Ergonomics, 32(6).

[4] Lariviere, C., Gagnon, D. and Loisel, P. (2002) A biomechanical comparison of lifting techniques between subjects with and without chronic low back pain during freestyle lifting and lowering tasks, Clinical Biomechanics, **17**, pp. 89-98.

[5] Randall, S. B. (2002) A Guide to Manual Materials Handling and Back Safety, In: Jeter, G. ed. Raleigh, NC: N.C. Department of Labor Occupational Safety and Health Program.

[6] Shaikh, I., Jayaram, U., Jayaram, S. and Palmer, C. (2004) Participatory Ergonomics Using VR Integrated With Analysis Tools, Proceedings of the 2004 Winter Simulation Conference.

Factors Affecting Event Detection in Dynamic Environment: the Case of ATC

Paola Amaldi
IDC, School of Computing Science,
Middlesex University
London
e-mail:p.amaldi-trillo@mdx.ac.uk

Jean-Marie Cellier
Université de Toulouse "Le Mirail"
Toulouse - France
e-mail cellier@univ-tlse2.fr

This work is part of a broader project intending to design a computer system to aid practitioners to detect events in a dynamic environment in the area of air traffic control (Leroux, 1997a, b)). We aim to provide further empirical evidence to the well documented phenomenon of perceptual chunking as a strategy to process a large number of information in a complex domain. We further claim that chunking is not only a way to parse events while monitoring a continuous stream of dynamic data, but also that problems are diagnosed by identifying these meaningful events which are claimed to be a basic unity of analysis in operational environments. Implications for designing aiding technology to event detection in operations environments are discussed.

Event detection, resource management, air traffic control

1. INTRODUCTION

In many high tempo environments, the role of perceiving (projected) changes of safety critical systems is never overstated. For example in Air Traffic Management (ATM) controllers have to monitor the process very rarely directly but more often through visual displays focusing on present and projected tendencies (for an overview of Air Traffic Control operations see Wickens, Movor, Parasuraman, McGee, 1998). Successful monitoring and problem solving involves the detection of (familiar) patterns as it has been largely discussed in the literature on expert performance (Chi, Feltovich, Glaser (1981); Chase and Simon, 1973) and naturalistic decision making. In highly dynamic environments, pattern perception involves the detection of meaningful changes over time, or to paraphrase Christoffersen and Woods (2003) the perception of *events* involving the integration of data over time. For example in the ATM domain a meaningful event is the detection of a (future) potential loss of standard separation between two or more aircraft. Detecting such conditions involves judgements based on assembling, comparing, remembering and projecting relevant data to arrive at the required decision. In many dynamic environments not all changes are events and in some cases lack of changes might be an indicator of a meaningful event. These judgements might be competing with other demands on attention and as such are vulnerable to interruption by alarms, by communication from other parties, by the need to attend to other parallel tasks. For example, one of the last minute factors that led to the en route collision over the lake of Constance (BFU, 2004) was that the only air traffic controllers left in charge of the entire Swiss Airspace, had to divide his attention between two concurrent tasks, the landing of an aircraft in a Swiss airport and the monitoring of the en-route traffic. The aggravating factor was that the two events were to be monitored on two different displays forcing the controller to hold two unrelated representations of events (for a comprehensive analysis of the accident, see the BFU Report)

2. MODEL OF CHANGE DETECTION

Christoffersen, Blike, & Woods (2003) have summarised their findings based on studies on domains such as practitioners NASA mission control and medical monitoring displays to propose a model for event identification which claims to include aspects and relationship among those aspects, responsible for the definition of events. Their model, which will be presented as an amended version based on our own findings, highlight the role of the observer's goals knowledge and expectations. This implies that facing the same situation, there might be different ways of "parsing" it into meaningful events, depending on the frame of mind of the observer(s). In other word depending on variables such as knowledge of the past, projection of future changes, models of the process control behaviour, different events might be detected by different observers while analysing the same situation. As for air traffic controllers, we expect that certain features of the situation will inevitably raise the attention of the observers, if not the cost would be failure to detect collision as in the tragic case of the en-route collision (BFU, 2004). Further the model implies that in a continuous streams of changes events are detected by noticing a discrepancy between expected changes and the actual or the projected. Expectations are the result of a model of what are the inputs that drive changes and how the system should respond to such inputs, in other words the system dynamics. For example the night of the TCAS accident, the Swiss controllers was reported to have issued twice to the Tupolev the instruction to descend flight level. The repetition was due to the latency of response by the Russian crew who was debating about whether to follow ATC or TCAS instructions. In this case knowledge of active influences and system dynamics led the controllers to identify a lack of change as a relevant event.

As part of a broader project, our study intended to investigate how air traffic controllers identified meaningful events by directing their attention to selected portions of the radar image. Following the model proposed by Christoffersen & Woods (2003) we aimed at identifying patterns detection strategies. The higher goal was to inform the design of an aid to identify "zones" where the intervention of controllers were desirable either to improve safety or traffic flow.

3. DESCRIPTION OF THE STUDY

Ten expert controllers from a French Air Traffic Control Centre were individually presented with a simulated air traffic scenario that was representative of a type of traffic most commonly encountered in the French airspace. The simulation was played back (no action control was possible) and was stopped at four pre-specified times. For each stopping time, controllers were individually interviewed and had to give a description of the general situation including identifying potential conflicts on the "frozen" radar image. The interviews were taped, transcribed and used later for analysis. For each stopping times, problems as described by the each controller were identified with the intent of

- verifying the consensus for each situation analysed by the ten controllers;
- identifying sources of individual variability
- validating and possibly refining the model of change detection developed by Chrostoffersen & Woods (2003)

4. RESULTS

For each of the four stopping time, the problems identified by air traffic controllers were mapped into a printout of the traffic image for that particular time. Four images were obtained but only two are reproduced here, for sake of space (the four images can be presented verbally). Before describing the two images, three general findings will be highlighted (for an extensive analysis see Amaldi, 1999).

The first observation was a high variability among controllers when identifying potential problems or potential loci to improve traffic flows. More precisely, out of 25 problems identified, only 3 were identified by at least 90% of the controllers interviewed. The order with which the problems were identified was also extremely variable with the exception of the 3 widely detected, there was hardly any pattern in the distribution of attention to detect the problems.

The second important result appearing on the mapping of the problems, is a clear partition of the radar image into "zones". Although, as for individual problems, the order with which the zones have been discussed by the controllers was highly variable, their definition was very stable, zones were not overlapping, meaning that controllers all use the same knowledge to partition the whole traffic situation.

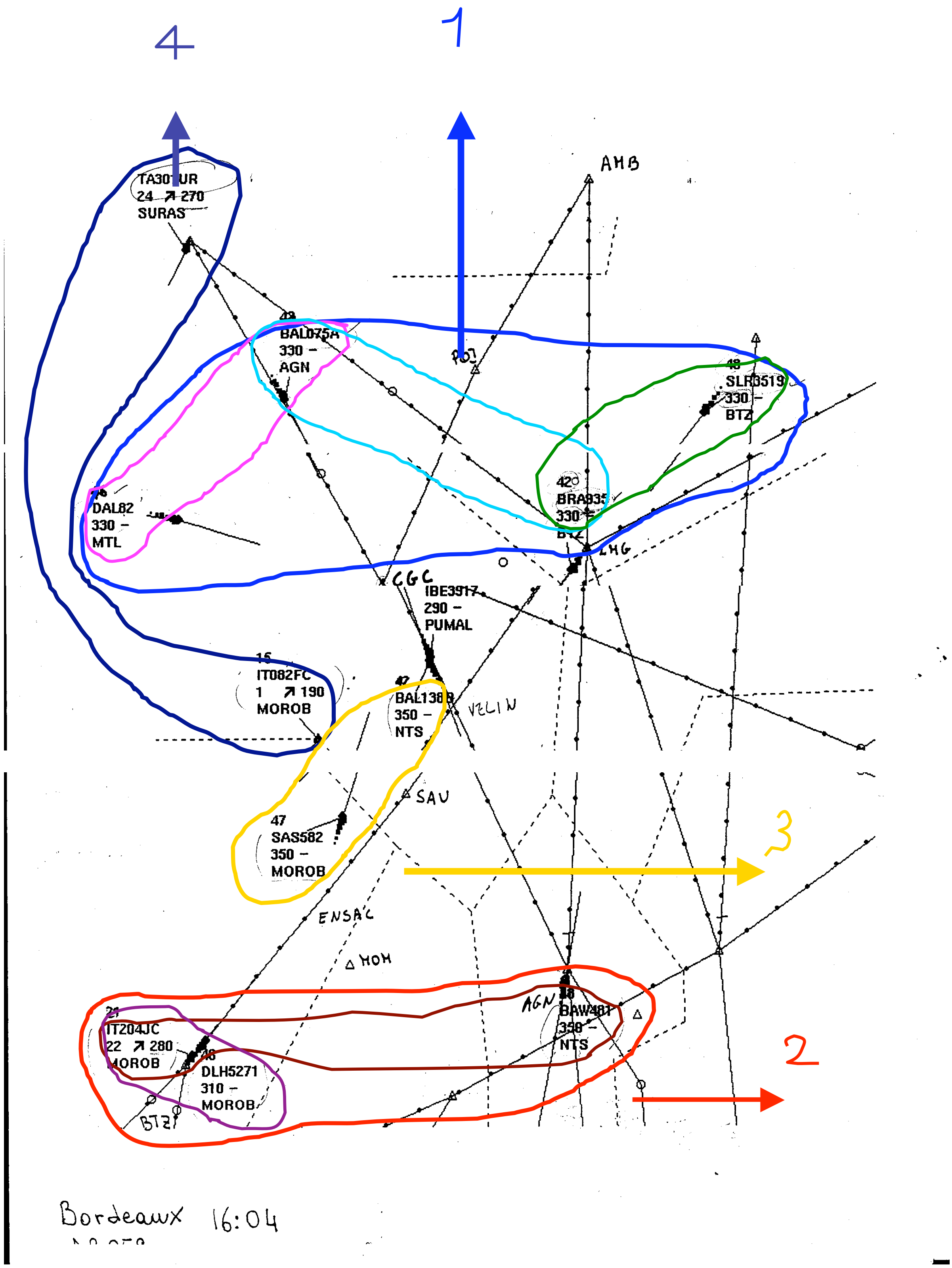


FIGURE 1: Graphical representation of “zones” identified by the controllers

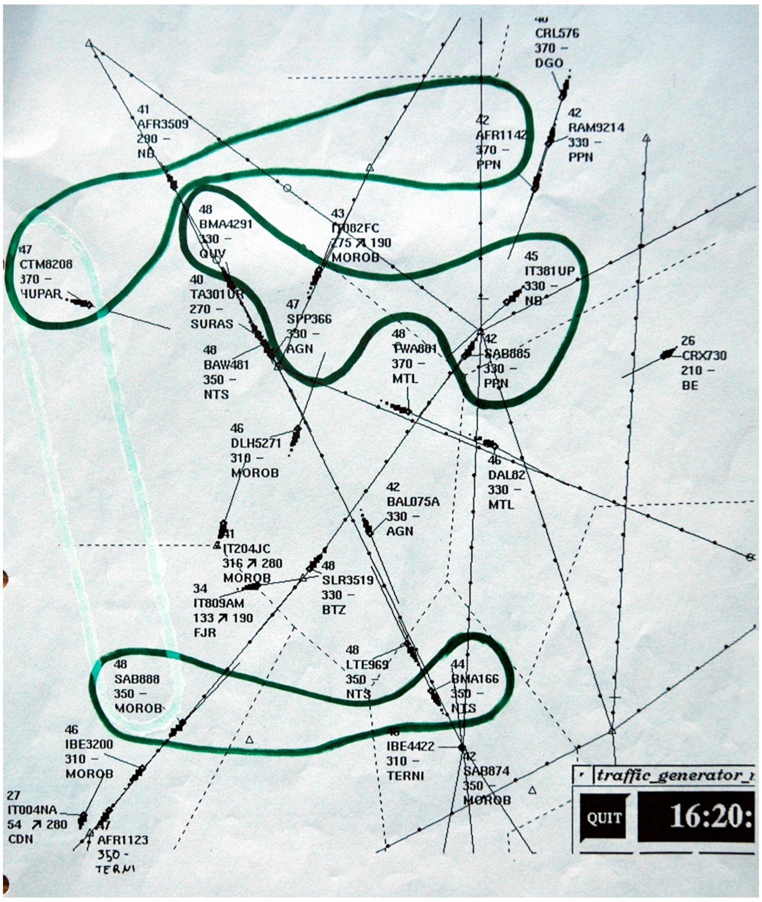


FIGURE 2: "Zones" of problems identified by controllers

Typically problems were grouped by same Flight Level (FL), same exit point from the sector, face-to-face climbing traffic, intersecting climbing trajectories, and excessive speed of traffic behind. However these knowledge factors alone would not explain the variability of the problem detection. Other factors will later be invoked.

Third, the two radar images reproduced here clearly show that the number and configuration of zones in the first radar image is more complex, i.e., more zones and more problems identified within each zone. And yet the traffic load in the first image is more than half than in the second image (12 vs 27 aircraft). As it has been pointed out by several studies on traffic complexity, our results confirm that the traffic load alone cannot be a reliable measure of perceived workload. Looking at figure 1, zone 1 is discussed unanimously by all controllers as it includes the detection of a potential collision between two aircraft.

Protocol's analysis has revealed potential sources of the observed variability. For example, the traffic configuration involving the IT204JC and the DLH5271 is regarded differently by a number of controllers. Some consider that the pilot's request to climb to FL 310 can be satisfied, others rule out such a possibility and make an early decision as to its final cruising level, and others consider that the problem is really to b effaced by the next adjacent sector. These differences in assessing the situation as revealed by the analysis of the protocols, points to an additional factors that was not considered in the original model of change detection as highlighted by Christoffersen & Woods (2003), *strategies for managing resources* (see figure 3). They include a perception of risk ("I am without constraint here because I can handle it."), perception of time available, temporal horizon, taking into account other parties' requests, co-operating with adjacent sectors. All of these factors, besides those already identified such as a mental model of the factors affecting the system behaviour, understanding the consequences of input into the system, have revealed to explain some of the observed variability.

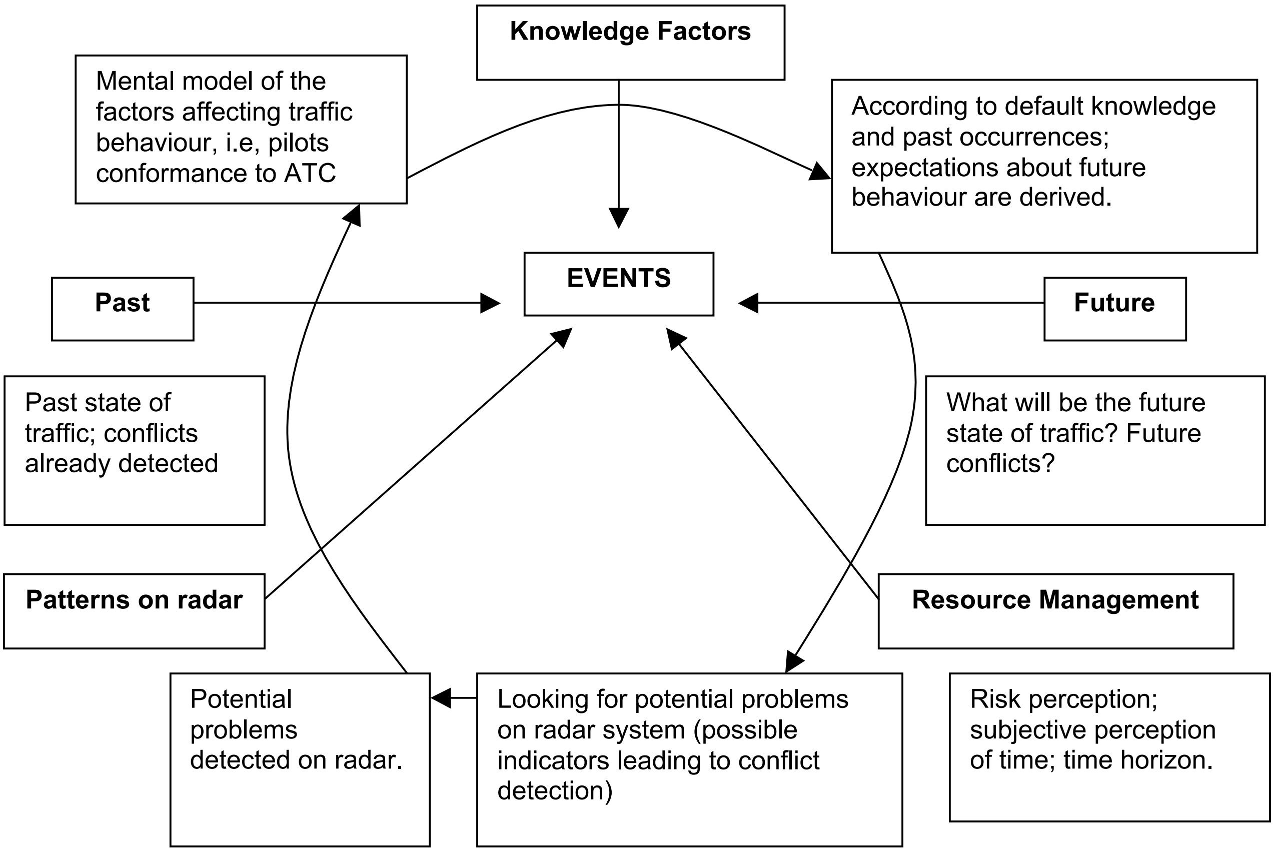


FIGURE 3: Framework for change detection in ATC (Adapted from Christoffersen & Woods, 2003)

To conclude, our results show on the one hand that the detection of meaningful changes or events, does occur by consistently partitioning the radar images into “zones”. Controllers seem to converge in recognizing the low level features and their salience in identifying common patterns of traffic. This could encourage the development of automated aid that would assist controllers in focusing their attention on the “relevant” information. However the high variability observed in the detection of problems suggests that the uncertainties of the development are weighted very differently. This in turn suggests that for any aiding tool to be found “engaging” and worth using, it has to allow for individual variability, as advocated by some literature.

REFERENCES

Amaldi. P. (1999). Le rôle d’un modèle du diagnostic dans l’évaluation d’un système d’aide à la décision dans le contrôle aérien: Le cas d’ERATO. Ph.D. Thesis, University of Toulouse “Le Mirail”, France

Bisseret, A. (1995). Psychologie cognitive de la décision. Représentation experte chez les aiguilleurs du ciel. Toulouse : Octarès.

Chase, G. W., Simon A. H. (1973) Perception in chess. In *Cognitive Psychology*, 4:55-81

Chi, M.T.H, Feltovich, P.J., Glaser, R (1981) Categorization and representation of physics problems by experts and novice. In *Cognitive Science*, **5**, pg. 121-152

BFU. (2004). German Federal Bureau of Aircraft Accidents *Investigation Report* available at http://www.bfu-web.de/cln_001/nn_53086/EN/Publications/Investigation_20Report/reports__node.html__nnn=true

Christoffersen, K., & Woods, D. (2003) Making sense of change: The challenge of events in operations environment. CSEL, The Ohio State University, USA. Available at http://csel.eng.ohio-state.edu/woods/CTA_Studies/MSoC2.1.pdf

Leroux, M. (1997a). Evaluation of advanced decision aids for en-route ATCOs: Towards an operational deployment. In Jensen, R.S., & L: Rakovan (Eds). Proceedings of the ninth International Symposium on Aviation Psychology, Columbus, Ohio, p.225-230

Leroux, M (1997b). Méthodologie de spécification d’interface guidée par un modèle cognitive. Application à la spécification des futures aides à la decision pour les contrôleurs aériens. In Huitièmes Journées de l’Interaction Homme-Machine. Toulouse Editions Cépaduès.

Wickens, C. D., Movor, A.S.., Parasuraman, R. McGee, J.P. (1998). The future of air traffic control: human operators and automation. Washington DC: National Academic Press

Emotional Response to System Messages – Are They Liked?

Sabine Wollstädter
University of Lüneburg,
Wilschenbrucher Weg 84,
21335 Lüneburg,
Germany
sw_78@web.de

Christian Peter
Fraunhofer Institute for Computer Graphics,
J. Jungius Str. 11,
18059 Rostock, Germany
cpeter@igd-r.fraunhofer.de

Hans-Rüdiger Pfister
University of Lüneburg,
Wilschenbrucher Weg 84,
21335 Lüneburg,
Germany
pfister@uni-lueneburg.de

Design aspects of user interfaces increasingly include emotional and hedonistic facets. With the steadily rising number of computer users with no technical background and no interest and not willing to understand what lies behind that shiny desktop, intuitive controls and information presentations are of growing importance. In this paper we report on the results of a study examining the emotional effect of system messages on users. Different ways to communicate the message were examined: pure text messages, spoken text, and figures, and combination of these with each other and with additional sound. Results show that system messages in general are not liked by users, regardless of their content and modalities used to convey them, and that spoken messages are those perceived as least annoying.

Emotion, system messages, affective response, physiology measurements, emotion detection

1. INTRODUCTION

Over the last few thousand years, the homo sapiens has developed very efficient ways to convey information. Different modalities are used for this, like sounds, signs, gestures, and language. They are used either separately, such as short utterances, gestures, or drawings, or combined, as we know from lively discussions, dance performances, or textual reports accompanied by pictures, sound and video. Humans even apply very fine-tuned modulations on their communications to convey additional information on e.g. urgency and reliability or subjective emotional evaluations of the message's content. Piercing eyes, a frown, or a subtle change in the tone of the voice are examples for these.

In the late 20th century however, humans programming computers seemed to be fairly unaware of the different ways to communicate with their fellow-humans, or at least they decided to let the computers start where their own ancestors started thousands of generations ago, with plain and mainly crude bursts of information. Computer applications usually use one and the same way to communicate all sorts of information, mostly in form of written text snippets (figure 1). As a result, people tend not to read text messages since they think they have read this one before and know which button to press in order to be able to carry on with what they intend to do.

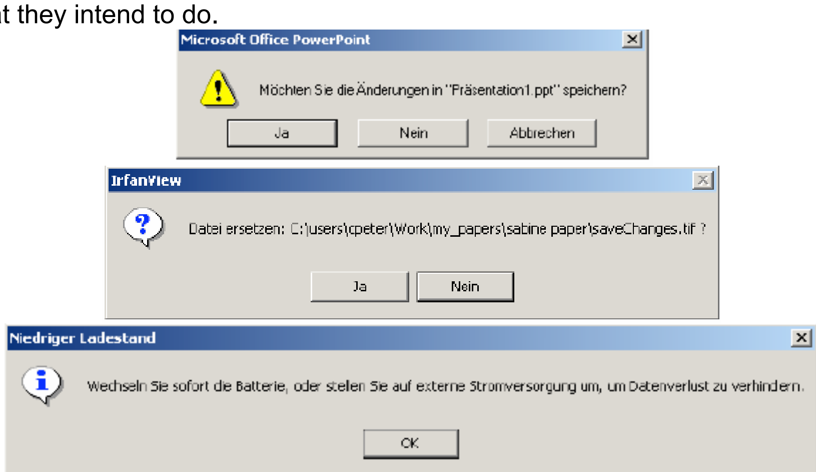


FIGURE 1: Save changes? (top) vs. Replace file? (centre) vs. Change battery now! (bottom). Note that the exclamation mark is not used to warn on anticipated total system failure by energy break down.

Smart programmers realising this added little graphics to the text messages. This really made people more attentive towards the presented message. Sadly, too many graphics with difficult interpretation have been tried so people were more concerned about the meaning of the graphics than the text. In a consequential step, the look of the used graphics has been simplified and their number has been reduced to about three for the operating system (depending on the system) and maybe one or two application specific ones. They were now thought to give additional and quick information on the basic nature of the message, e.g. to differentiate between notifications, warnings, and questions. Sadly again these are used in a seemingly random manner as can be seen in figure 1. In this example, the high risk information on low batteries has been equipped with an "i" for information, while "Do you want to save changes?" shows an exclamation mark, but "Overwrite file?" a question mark.

Trials went on using one more information channel: sounds. Very interesting studies have been performed here [12, 13, 15, 16, 18], with some of them having found their way into applications for plant monitoring, transportation systems, cooperative work environments, and even learning software. Most programmers however, and sadly even developers of operating systems, still neglect scientific findings and seem to see sound enhancements as a funny, meaning-free programming feature of pop-up messages. So sound enhancements quickly mutated to entertaining or annoying sound themes, decoupling the sound from the actual message and uselessly allocating system resources.

In this paper we investigate three ways used to communicate information to users, from the emotional point of view, i.e. how do users perceive those message pop-ups. Measuring physiological parameters correlated with human emotional reactions, we confronted test users with text messages, speech output, figures and sounds. Combinations of these were also used and the testers gave self-assessments of their emotional experience during the test. Our goal was to clarify if system messages communicated through different modalities elicit different emotional responses in the users.

In the next chapter we give an overview of the experiment performed, followed by a short description of the underlying emotion assessment strategies. Results are presented in the closing chapter.

2. THE EXPERIMENT

In the experiment, 54 students had to perform different (simple) tasks. These tasks were communicated to them using plain text (figure 2). Tasks included a typing task, a calculation task, and a search task. The typing task was a simple learn program for touch-typing. Users got feedback on typing errors they made, their progress in the job, and feedback on successfully performed lessons.

In the calculation task the students were given a bill in the look of a spreadsheet. They had to check the figures and correct them if they found any mistakes. Feedback was about non-editable fields in the spreadsheet, the negative effect of their changes to the overall sum, and question concerning saving the changes made.

The search task required the students to find a name in a list of names. There were several possibilities of the name being in the list, from not present, to just one, to many. The feedback was about no or how many names have been found, and suggested typing errors.

All scenarios had a neutral, a positive, and a negative (error) message. All messages were communicated by different modalities each, which were speech, figures (symbols), text, and additional sound, i. e. beeps. Those were used as single information channels, as well as in combination with each other. Combinations were text and symbol, symbol and speech, text and beep, symbol and beep, and speech and beep.

The overall time needed for one session was between 25 and 35 minutes.

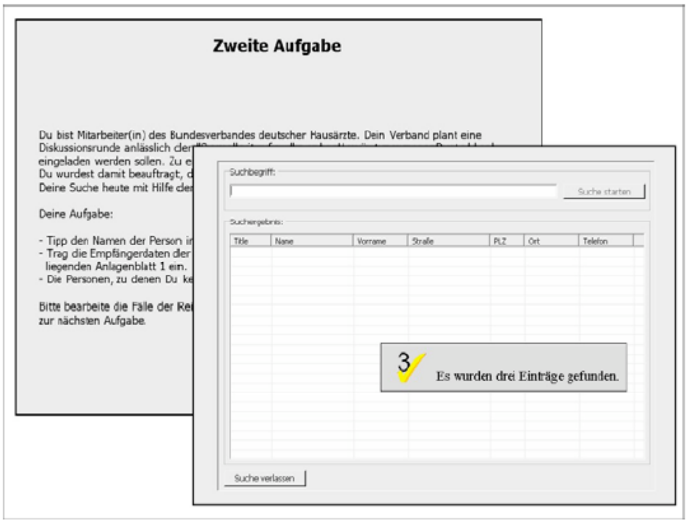


FIGURE 2: Example of task description, edit field, and graphical feedback for the search task.

The used test software has been developed by Fraunhofer IGD Rostock especially for configuring and setting up human-computer interaction experiments. It allows to easily create, add and change slides as seen in figure 2, to start and quit applications, and to log specific application and system events. Another benefit of this dedicated framework is that user interfaces can be easily designed with a look new to all subjects, which provides for same conditions concerning familiarity and pre-knowledge about the interface for all users.

3. EMOTION ASSESSMENT

Emotions manifest themselves in physiological changes controlled by the autonomous nervous system (ANS). Those changes can be easily observable ones, like mimics (facial expressions), gesture, or body movements, more subtle ones like changes in the voice, or modifications of physiological parameters usually not recognised by a human observer, for instance blood pressure, heart rate, or electro-dermal activity. Sensing and recognizing emotions with computers is a challenging but very promising undertaking, requiring the integration of hardware (sensors), mathematical methods for data enhancement and filtering, pattern recognition and classification. While facial expressions are one of the most obvious manifestations of emotions [7], detecting them is still a challenge (see [4]) although some progress has been made in recent years [1, 8]. Speech parameters have also been examined for correlations with emotions, with increasingly acceptable results [3, 5, 11]. Physiological indicators of emotions have been studied for a long time (e.g. [2, 6, 9, 17, 19] and can be considered as those best understood today. In this study we therefore decided for measuring physiological parameters using a skin resistance sensor, a skin temperature sensor, and a heart rate sensor. To correlate the physiological measurements with individual emotional experiences of the users, the subjects were asked to assess their subjectively felt emotion on SAM, the Self Assessment Manikin introduced 1980 by Lang [10], measuring the three affective dimensions valence, dominance, and arousal (figure 3).

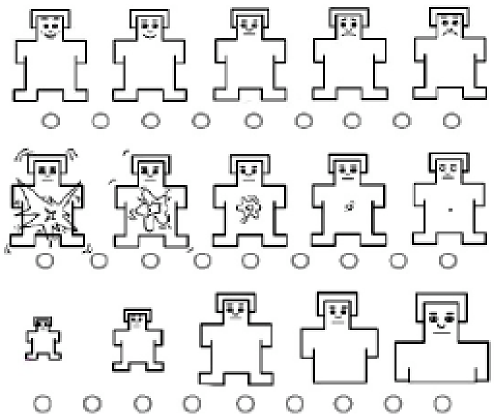


FIGURE 3: The Self Assessment Manikin (SAM) for valence (top), arousal (centre), and dominance (bottom), c.f. [14].

4. RESULTS

Analysis has been performed on the self assessment data, i.e. the SAM ratings, and on the physiological data. Self assessments were taken immediately after message events. Care has been taken not to harass the subjects with too many rating requests, the average interval between ratings was 100 seconds.

Physiology data show that the different communication channels result in very different physiological reactions of a user. Particularly spoken requests resulted in much higher physiological activation then the other three, particularly heart rate and skin temperature (table 1). Skin resistance has increased for any modality which might be due to the cognitive load associated with processing the messages. Concerning the message category, input requests showed stronger responses than the other two, but spoken error messages also showed more intense emotional reactions. Interestingly, the difference between message categories became insignificant with non-speech messages, text and figures alike. Looking at the effect of graphical symbols, we found that they have no significant impact on the physiological response to the shown message.

Analysis of the SAM ratings support the physiology findings. Spoken messages have been experienced much more pleasant than text messages, either with or without graphics and sound, respectively. Input requests resulted in more positive valence ratings than error messages or state notifications (neutral). Analysis of the SAM data also indicate that the subjective feeling of dominance in response to message events is stronger for spoken than for written messages. Dominance is felt strongest at input requests, medium at status messages, and smallest at error messages. There are also indications that system messages in general decrease positive feelings of users, but have no significant effect on a neutral or negative mood.

	Text	Speech	Input	Status	Error
Heart rate	+	++	+++	+	+
EDA	– no significant changes –				
Skin temp.	+	++	+++	+	+

TABLE1: Effects on physiological parameters

	Text	Speech	Input	Status	Error
Valence	++	+++	+++	+	+
Arousal	++	+	+	+	+++
Dominance	++	+++	+++	++	+

TABLE2: Effects on emotional experiences

To summarize, we found that the general hypothesis that users show substantially different and specific emotions in response to system messages could be largely confirmed. It could be shown that spoken messages generally elicit more positive and more dominant emotional response than do written messages. In general, any message is subjectively perceived as a negative or at best neutral event, regardless of modality and content. Not surprisingly it seems that users don't like to be interrupted by machines.

The findings suggest that system messages need to be designed more carefully, making sure users want to be informed and appreciate the system's desire to provide them with relevant information. Their phrasing and options provided should always give the users the feeling of being in control. Where appropriate, speech should be used to inform or communicate with the user.

REFERENCES

[1] Aleksic, P. S., & Katsaggelos, A. K. (2005). Automatic Facial Expression Recognition Using Facial Animation Parameters and Multi-Stream Hmms. IEEE Trans. on Sig. Proc. Supplement on Secure Media.
[2] Ax, A. (1953). The physiological differentiation between fear and anger in humans. In Psychosomatic Medicine 55 (5), pp. 433-442. The American Psychosomatic Society.
[3] Burkhardt, F., & Sendlmeier, W. F. (2000). Verification of Acoustical Correlates of Emotional Speech using Formant-Synthesis. In ISCA Workshop on Speech and Emotion.
[4] Cowie, R., Douglas-Cowie E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., & Taylor, J.G. (2001). Emotion recognition in human computer interfaces. IEEE Signal Processing Magazine, January.
[5] Douglas-Cowie, E. , Cowie, R., & Campbell, N. (2003). Speech and Emotion. Speech Communication, 40.
[6] Ekman, P., Levenson, R. W., & Friesen, W. (1983). Autonomic Nervous System Activity Distinguishes among Emotions. In Science, 221. The American Association for Advancement of Science.

[7] Ekman, P., & Davidson, R. J. (Eds.). (1994). The Nature of Emotion: Fundamental Questions. New York: Oxford University Press.
[8] Fasel B., Luettin J. (2003). Automatic Facial Expression Analysis - A Survey. Pattern Recognition, 36.
[9] Herbon A., Peter C., Markert L., van der Meer E., Voskamp J. (2005). Emotion Studies in HCI – a New Approach. Proc.2005 HCI International Conference, Las Vegas, Vol. 1, CD-ROM. ISBN 0-8058-5807-5.
[10] Lang, P. J. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications. In J. B. Sidowski, Johnson, J. H. & Williams, T. A. (Ed.), Technology in mental health care delivery systems (pp.119-137).
[11] Laukka, P., Juslin, P. N., & Bresin, R. (2005). A dimensional approach to vocal expression of emotion. Cognition and Emotion, 19 (5), Psychology Press, 633-653.
[12] Gaver, W. (1989). The SonicFinder: An Interface That Uses Auditory Icons, In Human Computer Interaction, 4(1), 1989, pp. 67-94.
[13] Gaver, W., Smith, R. & O'Shea, T. (1991). Effective sounds in complex systems: the ARKola simulation, Proc. CHI, 85-90.
[14] Morris J.D. (1995). SAM: The Self-Assessment Manikin – An Efficient Cross-Cultural Measurement Of Emotional Response. Journal of Advertising Research Nov/Dec. 1995.
[15] Müller–Tomfelde C., Steiner S. (2001). Audio-Enhanced Collaboration At An Interactive Electronic Whiteboard. Proc. 2001 Intl. Conf. on Auditory Display, Espoo, Finland, Jul 29-Aug 1, 2000, .pp.267-271.
[16] Mynatt, E., Back, M. & Want, R. (1998). Designing Audio Aura, Proc. CHI, Los Angeles, CA, 566-573.
[17] Palomba, D., & Stegagno, L. (1993). Physiology, Perceived Emotion and Memory: Responding to Film Sequences. In Birbaumer, N. & Öhman, A. (Eds.): The Structure of Emotion (pp. 158-168). Toronto.
[18] Partala, T. & Surakka, V. (2004). The effects of affective interventions in human-computer interaction. Interacting with Computers, 16 295-309.
[19] Picard, R.W., Vyzas E., & Healey J. (2001). Toward Machine Emotional Intelligence - Analysis of Affective Physiological State. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 23 No. 10.

Engaging Chronically Depressed Patients: Discourse Analysis and Clinical Information System Design

Barbara Mirel and Mark Ackerman
University of Michigan, Ann Arbor, Michigan U.S.A.

We present methods and preliminary findings from our field study of depression care management to show how ethnography and discourse analyses uncover processes essential to engaging patients in self-management. These processes are individualized and rarely articulated in production-scale information system (CIS) design. We discuss why they must be addressed and propose appropriate categories and traits for them in CIS that will not compromise the personalized quality of the care.

1. INTRODUCTION

Unlike health services for acute diseases, best practices for chronic care must help patients manage their illnesses for the course of their lifetimes. This self-management is crucial for reducing the soaring costs and devastating functional effects of nonfatal chronic diseases like depression [1]. Consequently, depression care management programs have been introduced as a layer of clinical care between patients and doctors to foster self-management, regularly monitor patients, and collaborate across services. Many patients in these programs may not go to psychotherapy but do get prescriptions from psychiatrists. For them, care managers may be the main source of “talk support” but this care management is not therapy. It is education for self-management and is perhaps the most challenging aspect of depression care to precisely capture in clinical information systems (CIS). This support is individualized – e.g. negotiating with patients *personally feasible* coping strategies – and varies in style, content, and delivery across patients. Because of this variability, many physicians are skeptical about capturing idiographic interventions in a CIS without diminishing them. Yet, if records of these interventions are absent in CIS, medical professionals have no evidence-based way to tie their effects to outcomes and empirically justify them as best practices. The linchpin of chronic care – fostering self-management - will remain a mystery.

In clinical trials, depression care management interventions have significantly improved patients’ health outcomes and remission rates compared to care as usual. Unfortunately, mental health specialists do not know exactly why. They admittedly understand very little about “the active ingredients of effective …complex interventions,” or “differential effectiveness … according to severity of illness” [2, 3149-3150]. At present, data about individualized self-management interventions are stored only in clinicians’ free text notes. To systematize these data, some CIS are beginning to include templates and natural language processing. But these approaches put the cart before the horse. They do not rest on well-studied, ecologically derived categories matched to the work clinicians do *as these expert clinical practitioners understand it* or use terms that *connotatively* convey to these experts the meaning and purpose of particular idiographic moves in patient-clinician encounters.

We propose initial categories and traits based on preliminary findings of an 18 month ethnographic and discourse analysis study of depression care management. Our focus is the “active ingredients” in interventions that affect patients in adopting self-management strategies to improve their health outcomes. We follow the many qualitative studies that examine what makes "invisible medical work" work [3, 4] as well as those that put ethnography and discourse analysis in the service of CIS design [5]. As in all ethnographies, we construct categories and terms from field data, in this case care manager interventions in patient encounters; we also distinctively frame them with an eye on incorporating them into CIS to guide practice [6]. Therefore, we create traits that, on the one hand, can be systematically and efficiently captured and, on the other, do not undermine the personalization and latitude required for clinical success. We conclude by arguing that assuring fit-to-purpose CIS designs for these interventions must be exploratory, evolutionary and an iterative process with actual end-users. As studies show, this design process is necessary for flexible systems that must be sensitive to the dynamic, variable quality of idiographic care and clinicians’ skepticism to capturing/reifying it [7,8].

2. WHAT ARE INDIVIDUALIZED INTERVENTIONS AND WHY DESIGN CIS FORM THEM?

Individualized interventions for self-management are best understood by comparing them to standardized care in an example from patient-care manager exchange. The following example of their coexistence and distinctiveness includes only the care manager portion due to consent constraints.

1 The last time we talked to you your depression was pretty moderately severe. If you're still feeling the same I'm pretty concerned.
2 Are you sleeping any better?
3 How's your appetite?
4 Not good? Have you felt like doing anything?
5 Not really?
6 So you spoke to your current psychiatrist?
7 Okay… who is your current psychiatrist?
…
8 See, what I'm concerned about …like you are…is your symptoms are the same.
9 It's frustrating isn't it?
10 I appreciate that, I do. I understand what you're saying and I don't want to push you into making a change you're not comfortable with. My only hope is …my concern is for you to feel better. .I think you'd feel most comfortable sticking with Dr. S and starting therapy next week.
11 That sounds good. That sounds like a very good idea.

Lines 1-7 above reflect standard monitoring interactions (fact-collecting). Following a scripted interview form, the care manager asks for facts about pre-specified issues like depression symptoms. Even in these pat sequences, however, the care manager does some personalization. She uses "we" and expresses concern in line 1 to create a relationship and to enhance the patient's receptivity later in the call to more overt idiographic interactions. In lines 8-11 the care manager individualizes more directly. Between lines 7 and 8, the patient complains of little sense of progress with his psychiatrist on medication issues. In 8 and 9, the care manager voices concern, commiserates with the patient, and indicates that the two of them are solving this problem together ("I'm concerned about…like you are"). In line 10, the care manager assumes the right and trust to offer nondirective and directive advice for how to take control of this frustration. In line 11 the care manager encourages and validates the patient's solution.

To analyze such exchanges and construct categories and indicators of individualized interventions, we unite several lines of research. Behavioral health change research and ethnographies of medicine and other service work establish that individualized (and often invisible) clinical processes play a determinant role in motivating and supporting patient self-management and in directing the trajectory of patient care [9-11]. Only idiographic interventions can build the trust and sense of belonging that shapes patients' cognitive appraisals in ways that motivate self-management [12]. Research in quality assurance argues that self-management interventions must appear in patients' continuity of care records [13]. Groundwork for categorizing these interventions has been laid by socio-linguistic and rhetorical studies of patient-clinician encounters. Due to the dynamics of patient lifeworlds meeting clinician medical worlds, categories must capture the multiplicity of potential meanings and possibilities for actions that exchanges convey. Ideally, as Vizenor argues, CIS traits and ontologies for these "inputs to care" and their meanings should capture *speech acts* of care, not *logical class*es of domain information [14]. Speech acts bring about results in the world and communicate meanings to participants that cannot be uncovered solely by logical or denotative interpretations. When these meanings get lost – as Vizenor argues they often do in healthcare systems - traces of consequential acts and related results in the world get lost. Our study strives to define alternative communicative/speech act categories for depression care management.

3. APPLYING ETHNOGRPAHIC DISCOURSE ANALYSIS TO CIS DESIGN

In constructing categories and traits for CIS, we drew on our ethnographic study of depression care management as background to frame our complementary discourse analysis of 20 patient-care manager encounters. Ethnographically, we studied a teleheath depression care management unit in a large medical center's Depression clinic. It is co-directed by a psychiatrist and primary care physician and employs five nurses and social workers as care managers. Over the 18 months of our fieldwork, the care management program grew from 60 to 735 patients, ranging in age from 19 to 92 - 75% women, and diverse ethnicities and races. We observed care managers make over 100 calls – 20 of which resulted in completed interventions. We conducted interviews with stakeholders, attended health team meetings, and analyzed relevant documents.

We audio- and video-taped 20 completed intervention encounters, transcribed them, and analyzed them using Cheng and Johnstone's methods of discourse analysis [15]. All but three transcripts capture just the care manager side of the conversation due to consent constraints. Care managers filled in what patients said after the call. Unfortunately, we could not trace health outcomes to observed interventions due to privacy issues. Our ethnographic findings suggest, however, that several recurrent problems (e.g. what interventions to use with “difficult patients” who monopolize resources) would have been better addressed had the health team had empirical evidence about “invisible” interventions. We are in the process of completing a content analysis to quantify the frequency of various topics and themes, their distributed occurrence, and correlations with interactional traits [16].

To derive categories and traits of idiographic interventions we analyzed commonalities and variations in discursive structures and content. We distinguished scripted from personalized interventions; identified regularities in personalization; and highlighted care managers’ different approaches to self-management. From these patterns we abstracted categories and indicators of “self-management work.” Following Vizenor’s caution to avoid distortions from logical classes of information, we drew on care managers’ own language and connotations when discussing calls and patients’ readiness for self-management.

4. RESULTS AND DISCUSSION

Commonalities and variability that shaped our definitions of categories and traits include the following:

> *Encounters are always a “hybrid discourse*.” That is, they embody both standard monitoring *and* idiographic support, with success in the latter depending on striking just the right balance.
>
> *Encounters follow interview forms* and are loosely structured into a sequence of (a) fact-based monitoring questions,(b) open-ended questions exploring patient events/concerns relevant to self-management, and (c) a second round of fact-based monitoring. This structure lets care managers control and avoid patients’ “invitations” to therapy but still engage them in joint problem solving to increase their self-awareness and coping strategies. Variation occurs in the extent to which care managers deflect patients’ elaborations and life stories, which seem to affect patient receptivity.
>
> *Encounters connect lifeworld and medical discourses in complex ways,* and care managers “read” these evolving subtleties to make ad hoc judgments about the subsequent selection of interventions.
>
> *Encounters have some combination of the following three processes of idiographic intervention* for fostering self-management: (1) developing interpersonal trust and connectedness (2) educating as needed to replace unproductive taken-for-granteds, and (3) engaging in joint problem solving to lower barriers and asymmetries between lay and medical worlds. Because the combined processes that care managers enact depends on how they differentiate patients, varied processes provide a window into clinicians’ assumptions about optimal “ingredients” for diverse patients. Assessing if these assumptions hold based on effects of interventions on outcomes is one reason to collect these data in CIS.

Based on these patterns, we defined five categories of data for capture and display in CIS. Two relate to patients’ stances – typically ignored in CIS but as intrinsic to the *interactive* dynamic of fostering self-management as care manager processes of intervention are. They are: (1) patients’ views of their disease. (e.g. depression being a quest, lost cause, matter of restitution); and (2) patients’ strategies for dealing with depression (e.g. making peace with trade-offs, “toughing it out,” consulting professionals, progressively chipping away at it). The other three categories involve care managers’ idiographic intervention processes: (3) supportive affect (e.g. expressing encouragement and interest or addressing fears); (4) bridges between medical- and life-worlds (offering medical remedies to lifestyle concerns, conjuring authority, surplus answers beyond what a question asks); and (5) joint problem solving (exploring candidate treatments, contextualizing experiences, discussing trade-offs and scenarios).

This set of categories and traits are evident in the following care manager discursive move:

> You know [she commented to the patient] there’s something to be said for tolerating [the unpleasant] side effects. And a lot of people make that trade off. I think that you’re smart to look at it that way. The functioning that you get in the rest of your life is very important.

Here, the care manager engages the patient’s stance (making peace with a trade-off) and intervenes by contextualizing the patient’s experiences, supporting/affirming the patient’s choice, and mixing lay and medical terms (respectively, “a lot of people” not “”patients” and “functioning”).

5. CONCLUSIONS AND IMPLICATIONS

Methods and findings from our 18 month field study of depression care management show how *in situ* user experience and discourse analyses uncover, at a detail level, processes essential to engaging patients in self-management. These processes are individualized and rarely articulated in production-level CIS design. Yet for widespread best practices, CIS must capture and display them, albeit in categories and traits that will not compromise the personalized care.

Such traits for individualized processes need to be defined in ways that cue clinicians' expertise. Traits like "toughing it out," for example, are inexact but connotatively speak volumes to clinicians about types of patients. The enterprise of naming idiographic interventions in meaningful/speech act ways is in its infancy. Therefore, categories and traits like ours need to be subjected to exploratory, contextual testing. In the next phase of our work, we intend to implement these traits in the CIS at our study site. Collaborating with IS specialists and clinicians, we will have care managers enter data about them during or immediately after a patient call; we will evaluate if these are the right interventions to capture to identify consequential effects and iteratively revise and modify them. Our goal is to establish a set of indicators that can be captured during or right after care management encounters to ultimately provide clinicians with valid, meaningful knowledge about the right interventions for the right patients.

REFERENCES

[1] Gredon, J. (2001) "The burden of recurrent depression: causes, consequences, and future prospects." *Journal of Clinical Psychiatry* 62 (suppl 2), 5-9

[2] Gilbody, S., Whitty, P. Grimshaw, J.,Thomas, R. (2003) "Educational and organizational interventions to improve the management of depression in primary care." *JAMA* 289 (23), 3145-3151

[3] Strauss, A., Fagerhaugh, S., Suczek, B., & Wiener, C. (1985). Social Organization of Medical Work. Chicago: University of Chicago.

[4] Heath, C., & Luff, P. (1996). Documents and Professional Practice: "Bad" Organizational Reasons for "Good" Clinical Records. Proceedings of the ACM Conference on Computer-Supported Cooperative Work (CSCW'96), 354-363.

[5] Berg, Marc and Els Gorman (1999) The contextual nature of medical information." International Journal of Medical Informatics 6 (1-2) 51-60

[6] Anderson, S., Hardstone, G., Procter, R., & Williams, R. (2007). Down in the (Data)base(ment) : Supporting Configuration in Organisational Information Systems. In M. S. Ackerman, C. A. Halverson, T. Erickson & W. A. Kellogg (Eds.), Resources, Co-Evolution, and Artifacts: Theory in CSCW (in press). New York: Springer

[7] Stiemerling, O., Kahler, H. and Wulf, V. (1997) "How to make software softer – designing tailorable applications." Proceedings of Designing Interactive Systems (DIS '97), Amsterdam, 365-376

[8] Mambrey, R, Mark, G., Pankoke-Babatz, U (1995)"Integrating user advocacy into participatory design." Proceedings of the Participatory Design Converence. Cambridge, MA, 251-260

[9] Bates, D., Kuperman, G., Wang, S., Gandhi, T., Kittler,A., Volk, L., Spurr, C., Khorasani, R., Tanasijevic, M., and Middleton, B. (2003) "Ten commandments for effective clinical decision support: making the practice of evidence medicine a reality." *JAMA* 10 (6), 523-530

[10] Cameron, L. D., and Leventhal, H. (Eds). (2003) *The self regulation of health and illness behavior.* New York: Routledge

[11] Gutek, Barbara (1999) "The social psychology of service interactions." *Journal of Social Issues* 55 (3) 603-617

[12] Hagerty, B. and R. Williams (1999). "The effects of sense of belonging, social support, conflict, and loneliness on depression." *Nursing Research* 48 (4), 215-219

[13] Institute of Medicine In: Page, A. (ed). (2004) K*eeping patients safe.* Washington, DC: The National Academies Press,

[14] Vizenor, L. (2004) Actions in health organizations: an ontological analysis. *Proc Medinfo* : 1403-1407

[15] Cheng, M and B. Johnstone (2002) Reasons for reason-giving in a public opinion survey." Argumentation 16, 401-420

[16] Kippendorff, Klaus (2004) *Content analysis.* 2nd Ed. Thousand Oaks, CA: SAGE

WebMGISQL 3D – Iterating the Design Process Passing Through a Usability Study

Vincenzo Del Fatto[1], Luca Paolino[1], Fabio Pittarello[2], Monica Sebillo[1], Giuliana Vitiello[1]
[1] Dipartimento di Matematica e Informatica, Università di Salerno, via Ponte don Melillo, Fisciano (SA) 84084, Italy
[2] Dipartimento di Informatica, Università Cà Foscari, via Torino 155, Mestre (VE) 30172, Italy
[1] *{vdelfatt, lpaolino, msebillo, gvitiello}@unisa.it,* [2] *pitt@unive.it*

A spatial query language, useful to querying spatial databases and retrieving spatial information, becomes efficient, effective and pleasant-to-use only if its surrounding environment is able to provide unskilled users with immediate and comprehensive access to the supplied visual operations. In this paper we describe how a usability evaluation process has been applied to improve the interface of the visual environment *WebMGISQL 3D*, a web application for querying and retrieving spatial information, with special focus on those phenomena where the third dimension is a relevant feature.

Visual languages, human-computer interaction, 3D environment, usability

1. INTRODUCTION

The main reason for using visual languages is that they are often far more convenient to the user than traditional textual languages. However, many times this convenience is hidden because the environment surrounding the visual language offers a low degree of usability, in terms of efficiency, effectiveness and users' satisfaction. This is also true for spatial query languages: their environment should be designed so as to provide unskilled users with immediate and comprehensive access to the supplied visual operations. In this field little has been done so far, even issue companies and organization operating with Geographical Information Systems (GIS) recognize that the usability issue is the major obstacle to wide exploitation of the potentials characterizing those tools.

WebMGISQL is a web application, conceived to support users in querying and retrieving spatial information related to those phenomena where the third dimension is a relevant feature [3,4]. A first step towards the development of a visual environment for manipulating spatial information was performed in [6], where the *Metaphor GIS Query Language* (*MGISQL*) was introduced.

The design process carried out for the web-3D environment, benefited from techniques related to usability engineering [1,2,5], including a consistent iterated prototyping process and a systemic approach to the usability evaluation of the different design phases. In particular, the usability evaluation activities involved in the different stages both expert evaluators and final users. In this paper we describe a part of the usability study we have conducted, in the stage when the first prototype of *WebMGISQL 3D* was available. The study consisted of think-aloud sessions involving two groups of sample users (skilled and unskilled about GIS and Database Management Systems - DBMS), whose results were collected and analyzed to draw hints on how to evolve the prototype towards a new more usable version.

This paper is organized as follows. Section 2 presents the methodology, scenarios, tasks, and procedure we followed during the usability test and results we obtained by the usability study. Section 3 shows how much of the interface was improved based on this simple user centred design process. The final prototype of the *WebMGISQL 3D* environment is depicted in Section 4.

2. SETTING UP THE USABILITY STUDY

WebMGISQL 3D allows users to interact with 3D scenes, manipulating visual objects (the *geometaphors*), which represent geographic data with descriptive and geometric components. The visual environment is meant to provide GIS users with any tool they may need to acquire information from the underlying geographic database. The usability study we describe in the paper, started from the prototype shown in Figure 1. Users could build queries by spatially arranging geometaphors in the central working area, selecting the topological, directional or metrical option. The 2D icons located on the left side (Layer and

Feature) are representative of object classes and single objects of the spatial database they are querying. These icons can be easily dragged into the 3D environment and successively moved around in order to compose a 3D spatial arrangement which is automatically translated into an standard SQL statement by taking into account the visual query interpretation (topological, directional, metrical).

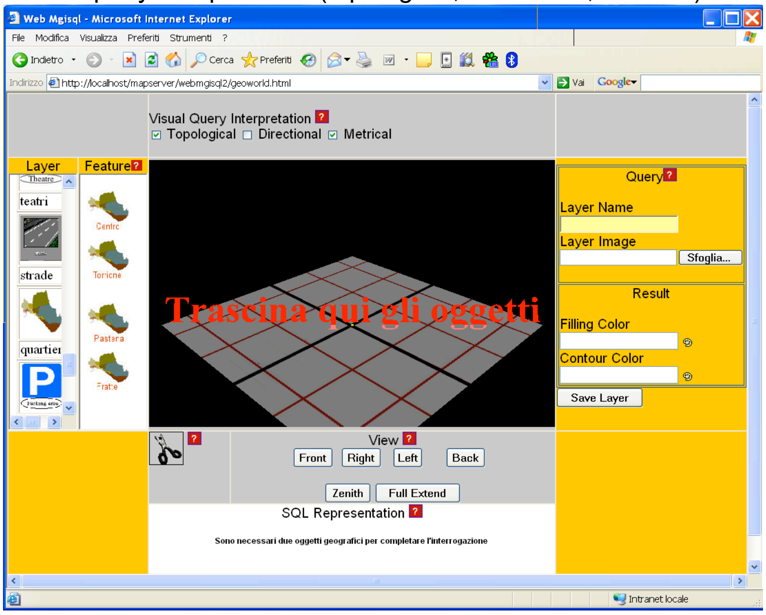


FIGURE 1: The *WebMGISQL 3D* prototype

As another relevant feature of prototype interface, while composing a visual query, some buttons allow user to change her point of view within the 3D scene, choosing among front, right, left, back, and zenith views. Finally, on the right area, a set of textfields allow user to specify information for storing and visualizing query results.

The prototype was submitted for evaluation to a group of users, with the goal of receiving feedback on how to improve the application. The subjects involved in the study were representative of one of the two profiles defined at the beginning of the design process, namely the *student* and the *database professor*. They were divided into two groups associated with different scenarios. Their contribution was very useful to lead to a refined version of the interface. The test included both quantitative and qualitative measures, in order to evaluate the system in terms of efficiency, efficacy and subjective satisfaction.

2.1. Methodology

Participants. We selected twenty participants, divided into two groups composed by 10 subjects each, who were totally unfamiliar with our application. Group A was composed by people corresponding to the student profile: they were males and females between the ages of 19 and 25, with no skill related to GIS/DBMS. Group B was composed by subjects belonging to the database professor profile; they were between the ages of 35 and 50 and had skills in the GIS/DBMS domain.

Apparatus. All tests occurred at the Department of Matematica e Informatica of University of Salerno. Both the tests were conducted in a quiet classroom after courses. A laptop computer working both as client and server was used. In order to record the interaction on the screen and the users' talking aloud, we used the Camtasia Studio™ software by Techsmith ™.

2.2. Scenarios and Tasks

Both the groups of the participants were asked to play scenarios compliant with those ones defined in Section 4.2, in order to allow the survey team to capture their performance and subjective satisfaction. Each scenario was described by a script that was given to the users:

> Script 1: Scenario A for the student profile.
> You play the part of a student going abroad for vacancy. You are a lover of eighteenth century paintings and you know that Salerno's churches are particularly full of this kind of pictures. You accommodate in a hotel located in the Mercatello quarter which is quite far from the historical centre of the city where the churches are concentrated. You have not much time and therefore you decide to look for churches within the historical centre; you decide also to rent a car for parking it within Concordia square because the hotel's receptionist told you that it is very close to the centre. You decide to take advantage of the WebMGISQL system, accessible from the hotel, for satisfying your information needs. After a brief self training you have to solve two different tasks, illustrated below.

Script 2: Scenario B for the database professor profile.
You play the part of a database professor teaching information systems and you want to explain the Spatial/SQL language taking advantage of the visual metaphors offered by WebMGISQL for permitting the users to understand more clearly the concepts of operands and spatial relations. In particular you have to compose visually two queries concerning the city of Salerno and involving different spatial operators; at the end you have to visualize the corresponding SQL code and show the results on a map. The specific tasks are described below.

While the groups performed their work in the context of different scenarios, we assigned the same tasks to both the groups, in order to have a wider sample for quantitative measures:

- *Task 1.* Look for all the roads located at Northwest of Concordia square and distant less than 1000 meters from the square.
- *Task 2.* Display the map showing all the churches within the historical centre quarter.

Both the tasks assigned allowed to test a number of significant functionalities of the prototype, including: selection and arrangement of geometaphors, selection of spatial relationships, SQL code generation and visualization of results.

2.3. Procedure

The survey team was present at the test and included a test assistant, an observer/note-taker and an observer/ time-keeper.

At the beginning of the session, the test assistant read to each participant one of the above scripts, according to the group the subject belonged to. Then, s/he illustrated the two tasks to be accomplished. Participants were not trained, but each subject was allowed to explore the functionalities of the interface by herself/himself before beginning the tasks. The observer/note-taker recorded the participants' actions, navigational errors, and verbal comments made during the task execution. The observer/time-keeper kept track of the time needed by each participant to complete the task and of her ability to perform it correctly. After the completion of all the tasks, participants were asked to answer to a brief questionnaire about the features of the interface, evidencing both the positive features and the drawbacks.

Test Measures. The test included both quantitative and qualitative measures. The former included time and number of completed tasks, the latter measured the subjective satisfaction of users. Errors made by the participants and logged by the survey team were used for making further design improvements to the interface, which led to the final version shown in Figure 2.

3. DISCUSSION ON RESULTS

The aim of our test was to verify the application potentialities on two main groups composed by novice users (group A) and GIS/DBMS experts (group B) and to validate the design choices that characterize the user interface.

The results obtained, partly summarized in Table 1, were encouraging. In particular:

- all the GIS/DBMS experts were able to complete the assigned tasks using the visual interface;
- the percentage decreased for novice users, varying from 75% to 87.5%. The latter result could be explained with the complete lack of knowledge about the GIS domain by the subjects belonging to group A;
- of course the knowledge of the domain also influenced the data on the time spent by the different groups for completing the tasks, as shown in Table 1(b). The difference is relevant for the first task: the novice users completed the task in 13'02'' while the GIS/DBMS experts needed only 4'35";
- the table shows also that the experience acquired performing the first task was a great help for diminishing the performance time during the execution of the second task. Such a result shows that the interface is not characterized by a steep learning curve and allows the users to perform significantly better even after the execution of a single task. In particular time results related to task 2 show that novice users obtain a meaningful improvement in the performance time, reducing from 13'02" to 8'56", while expert users reduce the execution time from 4'35" to 4' 04".

Tables 1(c) and 1(d), concerning the subjective satisfaction of users, revealed a sufficient appreciation of the overall interface and the visual composition of queries. The visual query composition got the highest satisfaction values, while the appreciation of the overall interface, although positive, was slightly lower due the identification of some interaction drawbacks, which were taken into account for the development of the final prototype. In both cases, novice users assigned higher satisfaction values, because the interface allowed them to overcome their inability to compose queries using alternative methods, such as SQL string composition.

	Group A	Group B
Task 1	87.5%	100%
Task 2	75%	100%

(a)

	Task 1	Task 2
Group A	13' 02"	8' 56"
Group B	4'35"	4' 04'

(b)

(scale from 1 to 5, easiest = 5)	
Group A	3.31
Group B	2.93

(c)

(scale from 1 to 5, easiest = 5)	
Group A	4
Group B	3.79

(d)

TABLE 1: (a) Percentage of participants who completed correctly the assigned tasks, (b) Time for completing the assigned tasks (the reported value includes also the time spent by the user on reading the task script, which is estimated in 30" for each task), (c) Overall interface easiness, (d) Visual query composition easiness.

4. THE FINAL INTERFACE

The observations during the task execution and the answers recorded after the test completion were analyzed and led us to develop a refined version of the user interface. Data collected during the usability study were analyzed in order to understand which points delayed the task completions, and which enhancements were needed to improve user's satisfaction. The resulting prototype is depicted in Figure 2. In particular, the following improvements were introduced:

- the map of the query results is now displayed on the right in a unified interface panel and it is updated each time the user presses the *Update Map* button, therefore the user is able to easily visualize query results.
- the set of widgets for storing the query results has been moved to a pop-up window which can be recalled after the visualization of the results. Then the user is no longer forced to specify the parameters for saving the query before data visualization, thus speeding up the process;
- the top of the interface displays the sequence of steps for posing a visual query, in order to help users to quickly formulate the queries. Each label appears on the top of the related interface widgets and it is associated to a short explanation. Such an adjustment was necessary because the correct step sequence was not always intuitive for the users;
- the SQL string located on the bottom side was replaced by a natural language query representation in order to improve readability of geometaphor arrangements. The SQL is still visible by clicking over the *show SQL code* button, added on the top of the interface;
- the set of checkboxes for choosing the spatial relationships was moved near the natural language query representation, in order to let users better understand the influence of spatial operators on the query composition;
- a new link, named *Get Started*, was added in order to help users to get acquainted with the application; such a link triggers an help page containing information about the application aim, the steps for performing the query and the meaning of the different interface objects;
- a *Delete* button was added to the interface for allowing users to delete the 3D scene. A similar widget was already available in the previous interface (the scissors icon), but its meaning was not intuitive;
- the labels associated to the selection of classes and instances of geometaphors (*layer* and *feature*) were replaced by descriptions (*category* and *object*) more familiar to the end users;
- finally, the *objects* repository was modified so as to display simultaneously all the instances. In the older prototype such a repository displayed only the instances corresponding to the class selected by the user in the *category* repository. Such a behaviour was not understood by some users that consequently were not able to find the objects they needed.

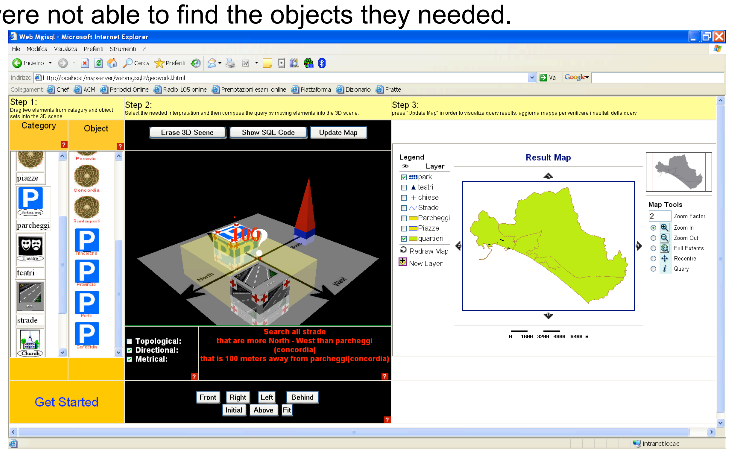


FIGURE 2: The refined version of the *WebMGISQL 3D* environment

REFERENCES

[1] Nielsen Jakob, (1994) Usability Engineering Book by, published by Morgan Kaufmann, San Francisco.
[2] Norman, D.A., Draper, S.W., (1986) User Centered System Design; New Perspectives on Human-Computer Interaction, Lawrence Erlbaum Associates, Inc. Mahwah, NJ, USA.
[3] L. Paolino, F. Pittarello, M. Sebillo, G. Tortora, G. Vitiello, (2002) “Extending the Metaphor GIS Query Language and Environment to 3D Domains”, *Proceedings of the Working Conference on Advanced Visual Interfaces - AVI 2002*, Trento (Italy), 22-24 May, pp. 144-147.
[4] L. Paolino, F. Pittarello, M. Sebillo, G. Tortora and G. Vitiello, (2003) “WebMGISQL - A 3D Visual Environment for GIS Querying”, *Proceedings of the 2003 International Conference on Visual Languages and Computing (VLC03)*, Miami (Florida), USA, 24-26 September.
[5] Shneiderman, B., (1998) Designing the user interface. Strategies for effective human-computer interaction. 3rd ed.,Addison-Wesley, Reading, MA.
[6] M. Sebillo, G. Tortora, G. Vitiello, (2000), “The Metaphor GIS Query Language”, *Journal of Visual Languages and Computing*, vol. 11, n. 4, pp. 439-454.

Guidelines for Designing Usable DVD Menus

Thomas Költringer, Martin Tomitsch, Karin Kappel, Thomas Grechenig
Research Group for Industrial Software (INSO), Vienna University of Technology
Wiedner Hauptstrasse 76/2/2, 1040 Vienna, Austria
{thomas.koeltringer, martin.tomitsch, karin.kappel, thomas.grechenig}@inso.tuwien.ac.at

Due to a lack of design standards, DVD menus often suffer from serious usability problems resulting in a negative impact on the user experience. Thus, the aim of this paper is to introduce new guidelines for designing usable DVD menus. This was accomplished through different usability-engineering methods focusing on expert walkthrough, user studies and card sorting. Finally, a prototype was developed as particularly useful tool for validating and improving the proposed guidelines.

Design methodology, user interface guidelines, DVD menus

1. INTRODUCTION

Usability inadequacies of DVD menus often cause frustration and thus have a negative impact on the user experience: a fact that gains more importance nowadays as the number of DVDs on the market is increasing. Users expect to have a good experience (quick and easy to operate, skill transfer between DVDs should be possible, nice introduction to the movie) when handling the DVD menu. However, DVDs often feature unintuitive and cluttered menus that make navigation difficult. Furthermore, inconsistencies between different DVD menus make skill transfer impossible. The main reason for usability problems and inconsistencies in DVD menu control is the lack of design standards: they simply do not exist [7], comparable to the early years of the Internet, when web pages were designed without following any standards or User Interface (UI) guidelines. Hence there is a strong need for developing DVD-specific UI guidelines to assure a certain level of quality for DVD menus.

UI Guidelines can be used for designing a UI as well as for evaluating an existing application in terms of usability. Designing UIs is a difficult and time-consuming task [6]. To provide the basis for high-quality UIs, the skills and the experience of the designer are key elements in addition to theories and guidelines [6]. Accordingly, Tetzlaff and Schwartz identified shortcomings and disadvantages of guidelines [9]. However, as will be shown below, well-developed UI guidelines provide reference and guidance during the design process [2]. Furthermore, the adoption of guidelines can reduce the number of iterations of the interface development process [2].

2. MOTIVATION

Designers of DVD menus have to face many DVD-specific design challenges, concerning usability issues as well as the technological implementation issues. DVD menus are different from traditional computer interfaces, such as software products or web pages. Furthermore, DVD users do not necessarily have any computer experience. They might not be familiar with traditional concepts, such as menus or buttons. In addition, user interaction with DVD menus is very different to interacting with software or a web page (in case the DVD is viewed on a stand-alone player or console). Designers also have to keep in mind that remote controls typically provide a 4-way navigation with special buttons (menu, submenu, language, subtitles, angle, etc.) instead of a pointing device.

3. DEVELOPING THE GUIDELINES

The development of UI guidelines is a process requiring experience as well as extensive application of various usability-engineering methods. To establish applicable guidelines that improve the quality of the product effectively, it is necessary to include different stakeholders (users, designers, developers, etc.) in the development process. This was accomplished through defining a framework consisting of three main phases. Whereas details concerning the first phase and insights into the development process have been published previously [4, 5], this report outlines applied methods and results gained in the course of the different phases.

Literature Review, Expert Walkthrough and Initial User Study

The first task to build a solid foundation for guidelines was to get an overview of significant literature published on the topic. This *literature review* clearly showed that research into this field is essential, because to date no scientific literature concerning DVD menu design is available. However, beside articles concerning technical aspects of DVD menu design, we found some publications on the web [1, 7]. To gain more information, we reviewed discussion forums and weblog entries discussing usability issues of DVD menus.

As a next step, an *expert walkthrough* was carried out to identify the main problems of navigating through DVD menus and to identify standards such as structure or wording. This was achieved by inviting three usability experts to investigate 70 different DVDs. In this context, 85 issues were revealed, ranging from general consistency problems to very specific solutions. On the one hand, issues concerning navigation proved to be essential, on the other hand animations and background music clearly addressed the user experience.

To reflect the knowledge and experience of users, we performed a user study collecting both quantitative and qualitative data. Hence, we decided to conduct a usability test and a survey. For the *usability test* we recruited 20 subjects (6 female, 14 male) with an average age of 22.5 years (SD=2.46). Four subjects watched DVDs weekly (experts), 9 subjects watched monthly (intermediate users) and 7 were novice users with little or no experience (light users). We selected 10 movie DVDs with different types of DVD menu implementations. The test consisted of five tasks that all subjects had to carry out (e.g. select English subtitles; select director's commentary; select a certain chapter). These tasks covered the typical usage pattern of DVDs. Expert users did not experience any serious problems, mainly because of better experience and good shortcut knowledge. In contrast, major problems arose for the novice and intermediate users when accomplishing the tasks. In the majority of cases, inconsistencies concerning the scene selection between the different DVDs were the crucial factors for handling problems. The subjects expected the navigation through the scenes to be sequential, which was not always the case. Furthermore, the mapping of the remote control to the menu navigation was not clear to most of the users. Another problem was the distribution of menu entries on several pages, e.g. the separation of languages on two pages in the language selection. Moreover, arbitrarily arranged menu items posed a problem for the subjects. During the post-test interviews the subjects expressed their annoyance about long 'unskippable' intro sequences and transitions between menus. Especially the scene selection menus often include long and annoying transitions. Overall, the user test resulted in 43 issues.

For the *survey* we set up an online questionnaire to collect solutions for specific questions of interest, such as wording, menu structure preferences, etc. We included open questions as well to obtain opinions and ideas for improvement from users. The target group for the questionnaire consisted only of experienced users. Hence, we invited users of several DVD portal sites and mailing lists to take part in our survey. We received 350 replies (23% female, 77% male) with an average age of 28.16 years (SD=7.76) from the questionnaire. Sixty-four percent watched DVDs at least once a week (experienced users), and 46% were owners of 25 or more DVDs. We analysed the data from the questionnaires to obtain usable and comprehensible solutions for wording, alignment of menu items, number of scenes, preferred bonus material, etc. We found that items in the main menu should be limited to *play movie, language selection, subtitles, scene selection*, and *bonus material*. Furthermore, the results concerning the scene selection

Short description	Arrange scenes either in a line or in a grid.
Detailed description	The scene previews should be aligned either in a line (horizontal or vertical) or in a grid. This ensures better mapping between the navigation buttons on the remote control and the scene layout.
Example	Six scenes are displayed on one page of the scene selection menu. They are arranged in a grid of three by two.
Priority	★★★★☆ (4/5)
Origin	Expert walkthrough, usability tests
Reference	3.2 (Menu Alignment), 4.1 (Clear Mapping), 7.7 (Grouping of Scenes)

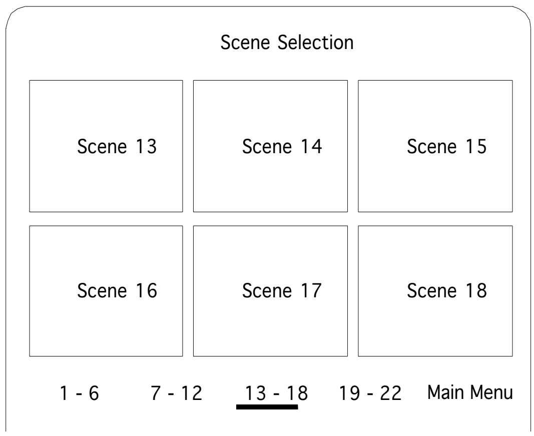


FIGURE 1: Guideline and illustration for the arrangement of scene previews (example)

showed that a movie of standard length should be divided into approximately 20 scenes. At least four, but at most six scenes should be grouped on one page. From our point of view, a very surprising result was the

Introduction
Keep the entire intro sequence short
Make all intro sequences skippable
Provide a menu for DVD menu language selection and possibly aspect ratio
Main Menu
Constrain the main menu to the following items: Play Movie, Language Selection, Subtitles, Scene Selection, and Bonus Material
Use the following labels for common items: Play Movie, Language Selection, Subtitles, Scene Selection, and Bonus Material
Put the initial focus of the main menu on Play Movie
Place the Play Movie item first (left topmost)
Include the movie title or other characteristic elements in the background picture of the main menu
Avoid non-descriptive text on a menu's background picture
Menus
Provide a maximum of only three levels of hierarchy
Align all menu items in a line or in a grid
If a menu allows the selection of one or more options, make all selections visible to the user
Show the selected option for menu entries where multiple option selection is possible
Set the focus on the corresponding menu item after returning from a submenu
Use a menu design that matches the movie
Try to group all menu items on one single page. If not possible, indicate clearly that there is another page for that menu
Do not animate menu items
Navigation
Use a clear mapping between the menu layout and the navigation buttons
Allow cycling through all menu items with one button
Make every navigation operation reversible
Provide transitions between menus, but make them skippable
Make the transitions short

Movie
Support the Menu button during movie playback
Support 'short-cut' buttons during movie playback
Allow resuming the movie via the Play Movie item
Settings
Provide either two separate menus for language and subtitles selection or combine them both in one menu
Keep all subtitles on one page
Do not automatically resume the movie after changing a setting in the menu
Scene Selection
Divide a movie of standard length into approximately 20 scenes and provide significant entrance points
Provide a scene index for quick page navigation
Highlight the currently active scene block in the scene index
Arrange scenes either in a line or in a grid
Allow sequential navigation trough all scene previews
Support cycling through all scene previews
If more than six scenes are provided on a DVD, split them into groups of four to six
Do not display any other menu items or elements between the scene previews on one page
Provide characteristic pictures or videos for the scenes in the scene selection menu, as well as the number of the scene
Avoid transitions between pages in the scene menu or make them skippable
Bonus
Provide bonus material, such as making of, deleted scenes, interviews, audio commentaries, and trailer(s)
Place the audio commentaries into the bonus material menu
Provide background music
Avoid distracting background music
Make background music tracks long enough
Use only background movies if they are not distracting

TABLE 1: The proposed guidelines for DVD menu design (short descriptions only)

fact that customers rarely consume bonus material, although it seems to be important for the purchase decision. In sum, the questionnaire delivered 12 new issues.

Building and Sorting the Guidelines

After sorting and reviewing all issues, we derived 44 guidelines for the design of DVD menus (see Table 1 for the short description of the guidelines). The detailed compilation of our guidelines was published in [10]. Each guideline was confirmed by at least two different methods used in the first phase. We also added a detailed description, an example scenario, and a priority as well as a statement about the origin, in order to receive simple, meaningful and descriptive guidelines. The priority of the results obtained from the usability tests was assigned according to the number of subjects who encountered the corresponding usability problem. Figure 1 shows an example for a DVD-specific guideline. For some guidelines we also added illustrations for better comprehensibility. We structured the guidelines by conducting *card sorting* [8]. We involved 5 male participants (computer science students) with an average age of 23 years (SD=1.58). The card sorting session resulted in the following eight categories: *Introduction, Main Menu, Menus, Navigation*, *Movie, Settings, Scene Selection,* and *Bonus.*

Validating and Improving the Guidelines

To evaluate and improve the elaborated guidelines, we handed the guidelines to a DVD menu designer, who implemented a working *prototype*. This was an important step since different users (designers, evaluators and users of the actual system) should benefit from the guidelines. The prototype allowed us to *test the guidelines* in terms of usability, readability, consistency and applicability.

Furthermore, we asked the designer to give us feedback on the technical feasibility of the guidelines. During the development of the prototype we found out that the realisation of some features causes severe technical problems. For example, showing the respective current settings in the language selection menu, is not possible, because the number of menus is limited to 10,000, and every combination of language/subtitles represents a single menu. Moreover, acoustic feedback represents a challenge since it is only possible to play one audio track per menu. Hence the developer can decide to either use this track for one continuous audio background loop or to provide audio feedback. There are many more technical limitations to face when implementing a DVD menu.

Finally, the prototype was subjected to a *usability test* to ensure that the application of the guidelines leads to an improvement in usability. We used the same test setting as in the usability test during the first phase, but extended the number of tasks to nine. We tested 12 (3 female, 9 male) participants (3 light, 7 intermediate, 2 expert users) with an average age of 24.92 years (SD=4.93). Only 2 intermediate and light users experienced some minor problems while accomplishing the tasks. Two less experienced users tried to find the audio commentary in the language settings, which was actually located in the bonus material menu. However, the other users did not hesitate and immediately selected the bonus material. Three users experienced problems in the scene selection. This menu included an index for quick navigation, which was not directly associated with the pictures of the scenes above. This result shows that it is necessary to attach the numbers of the scenes to the pictures. Beside this feedback, the participants did not face any problems while navigating through the menu.

We finally verified the quality of the guidelines in terms of comprehensibility and applicability. Therefore we handed the entire set of guidelines out to 10 groups of 4 students each. Their task was to evaluate 30 different common movie DVDs using *heuristic evaluation*. We received a lot of feedback regarding the guidelines and their structure. Most of the feedback concerned providing better navigation through the document (e.g. using an index or references) and offering more general information (e.g. definition of terms and remote control features). Therefore, we added references and definitions of DVD specific terms (e.g. remote control features) to the guidelines. Further feedback concerned the selection of language and subtitles in one or two submenus and the adjusting of wording.

4. DISCUSSION AND CONCLUSION

It is commonly accepted that guidelines are helpful tools when designing user interfaces [2]. Guidelines can also improve the value of a product for the customer. For example, the success of a shopping website depends largely on its usability. If customers have problems to select a product or purchase goods, they will leave the site without buying anything. A major difference to consumers of DVDs is the fact that customers buy the product without knowing the usability of the DVD menu (which is the same for mobile phones or video games). Furthermore, the menu is not the determining factor for purchasing a DVD. This might be a reason why production companies do not focus on usability issues of DVD menus. Nevertheless, usability inadequacies of DVD menus can frustrate customers and hence influence their future purchasing behaviour.

Other reasons for inadequate design of DVD menus are technical constraints discussed in this paper, and the fact that production companies try to minimise production costs.

Our study clearly showed that a careful design of DVD menus is essential for avoiding frustration of customers. The set of guidelines presented in this paper aims at providing a basis for developing usable DVD menus. However, many questions remain unanswered. Future research work will focus especially on content specific guidelines for DVDs (e.g. what is special for series, music DVDs or box sets). It is important to state that in general guidelines are a subject for continuous improvement especially in the fast changing world of electronic media.

5. ACKNOWLEDGEMENTS

We would like to thank Marc Gyarmaty, Markus Gärtner, Bernd Moser, Krunoslav Barta, Mohammad Tabatabai Irani and the students from our usability engineering class in spring 2005.

REFERENCES

[1] De Lancie, P. (2000) *Post Linear: Menu Design for DVD*. http://millimeter.com/ar/video_post_linear_menu/ (last date of access: May 4, 2006).
[2] Henninger, S., Haynes, K. and Reith, M.W. (1995) A Framework for Developing Experience-Based Usability Guidelines. *Proceedings of DIS 1995*, pp 43-53, ACM Press.
[3] Henninger, S. (2002) Tools Supporting the Delivery and Application of Usability Patterns. Patterns Workshop. CHI 2002.
[4] Kappel, K., Tomitsch, M., Költringer, T. and Grechenig, T. (2005) Developing User Interface Guidelines for DVD Menus. *Extended Abstracts CHI 2006*, pp 177-182, ACM Press.
[5] Költringer, T., Tomitsch, M., Kappel, K., Kalbeck, D. and Grechenig, T. (2005) Implications for Designing the User Experience of DVD Menus. *Extended Abstracts CHI 2005*, pp 1565-1568, ACM Press.
[6] Myers, B.A, (1993) Why are Human-Computer Interfaces Difficult to Design and Implement? Carnegie Mellon University, CMU-CS-93-183, Pittsburgh, PA.
[7] Norman, D.A. (2001) *DVD Menu Design: The Failures of Web Design Recreated Yet Again*. http://www.useit.com/alertbox/20011209.html (last date of access: May 2, 2006).

[8] Robertson, J. (2001) *Information design using card sorting*. http://www.steptwo.com.au/papers/cardsorting (last date of access: May 2, 2006).
[9] Tetzlaff, L. and Schwartz, D.R. (1991) The use of guidelines in interface design. *Proceedings of CHI 1991*, pp 329-333, ACM Press.
[10] Tomitsch, M., Költringer, T., Kappel, K. and Grechenig, T. (2006). DVD Menu User Interface Guidelines. Research Industrial Software Engineering (RISE), RISE-01/06, Schwechat, Austria.

Using RFID, Bar-code and Speech Technology to Support Context-Aware Maintenance with Wearable Computers

Chris Baber, Anthony Schwirtz, Tracy Hartley*, Tony Cousins*
EECE, The University of Birmingham, Birmingham
* Aerosystems International, Yeovil
c.baber@bham.ac.uk

In this paper we report the development of a prototype system that supports hands-free maintenance through recognition of the tools that a person is using and the components that they are working on. The concept is intended to use ubiquitous computing as a means of supporting direct interaction on objects in the world whilst supporting indirect interaction with a digital representation of that world (in this case, the diagrams and procedures used to support maintenance tasks). By simply tagging the tools (using RfiD) and providing a means of recognising bar-codes on components, it is possible to guide a person through a maintenance schedule. Speech technology is used to both display the content of the instructions and to manage this content. Video capture is supported to allow procedures to be captured and submitted to the manual authoring process (in order to simplify the complex and time-consuming process of setting up appropriate video recording for all stages of the procedure).

Wearable computers, maintenance, context-awareness, speech technology, RFID

1. MAINTENANCE AND WEARABLE COMPUTERS

It has long been recognised that wearable computers can support maintenance activity. Preliminary work, by Carnegie-Mellon University (CMU) in 1995 [1], developed the concept and applications in industrial settings followed soon after (most notably by Boeing [2]). There has been some work on the interaction between person and computer, e.g., CMU developed a novel interaction mechanism in their VuMan projects, which involved using a rotary dial to tab between elements in the display [1]. However, this involves stopping the task 'in-the-world' in order to interact with the computer. It is, of course, a moot point at to whether such a form of interaction interferes with activity in the world, or whether this interaction occurs at natural 'break-points' in the task. In this project, we began with the assumption that the maintenance engineer's focus should be on the components and tools in the world. This follows the notion of, for example, Nilsson et al. [3], in their study of ubiquitous computing for waste-water plant operators, speak of the need "...for systems that support the process of physical inspection and...to make the transition between physical interaction with the plant and interaction with digital representations of the plant smooth." (p.31). In a similar vein, we felt that if it was possible to signal which tool was being used, then it should be possible to determine which procedure step is being performed or whether the tool is appropriate to the current procedure step. Furthermore, if one can determine which tools are being used then it might also be possible to obtain data directly from some of the tools, e.g., readings from digital torque wrenches or callipers. Thus, the starting assumption for this project was the support of maintenance work as a form of ubiquitous computing in which interaction with the tools and components supported interaction with the digital representation of the aircraft being repaired or maintained.

2. BACKGROUND TECHNOLOGIES

Many industries employ electronic documents to support procedures. This is particularly prevalent in the aviation industry [4]. The documents provide engineers with procedures, diagrams, additional information that is (typically) in the form of hyperlinked documents. A subset of this material is the job card. This contains the task steps and diagrams necessary to complete a particular maintenance procedure on a particular component. In this work, we use the Aerosystems International *E-card* application. This is an electronic work card that presents information to the user (typically on a laptop computer) in the form of HTML pages that can be navigated in an Internet Explorer pane. The pages are automatically extracted by Aerosystems International's *Sapphire* electronic document system.

2.1 Bar codes and RFID tags

The aviation industry, in particular, is expanding its efforts to mark individual components. For large components, the trend remains the use of bar-codes. These are two-dimensional marks printed on labels that can be fixed to, or etched onto components with no interference to performance. The labels can be read using a bar-code reader. The U.S. Army is rolling out a programme to have Unique Identification (UID) of all material that it manages. Much of this UID will involve bar-coding of components (although RFiD is also possible). This would mean that, when working on a component, the maintenance engineer could read its code (using a bar-code reader). Having the bar-code reader mounted on the person, perhaps on the wrist or forearm, would allow line-of-sight scanning of bar-codes, and the scan to be read into a processor on the person [5]. Such a system could be used for stock checking. In a maintenance application, scanning of the bar-code identifies the part being worked on, and can be used to confirm that the engineer is working on the correct part. Recently, organisations have begun to use Radio Frequency Identification (RFiD) tags in preference to bar-codes. These are conceptual similar to bar-codes (in that there are read using a special reader and contain codes identifying the product), but do not rely on line-of-sight for reading. This means that it is possible, for instance, to embed a tag in the handle of a tool and for the tag to be read when the tool is picked up (by an antenna fitted into the cuff of the user's jacket). For this project, we make the assumption that components are bar-coded, and tools are RFiD tagged. Of particular, interest to the work reported in this short paper, is the manner in which simply picking up tools or scanning components' bar-codes can be used as a means of interacting with a computer.

2.2 Speech technology

Obviously, wearable computers are intended to support mobile interaction, and this is analogous to the aims of speech technology, e.g., hands and eyes free interaction. In our application, we use speech synthesis to read out the content of job cards (see below). This offers the potential to allow the engineer to be informed of the procedure step (and any warnings) whilst looking at the component. In order to keep the system fully hands-free, we use speech input to perform simple document navigation commands, e.g., <next>, <back> or procedure commands <details>, <confirm> etc.

2.3 Head-mounted displays and webcams

Wearable computers typically employ monocular, head-mounted displays (HMDs). This is primarily a question of cost, but also relates to the assumption that the wearer would be able to switch attention between the HMD and the world [8]. The HMDs we used had a fairly low resolution, i.e., 640 x 480. This is not unusual in monocular HMDs, but did require some redesign of the E-card pages so that their content could be displayed in a readable manner. Electronic manuals can present video to the user, e.g., to illustrate particular maintenance activities. A significant overhead associated with the production of such material relates to the capture of video; if one sets up a particular maintenance task for video capture, then this requires access to equipment and engineers that could prove costly. We assumed that it would be useful to have maintenance engineers video activity as they were doing their work. The videos could then (at the engineer's discretion) be submitted to the Manual Production Authority (MPA), which could include the clips in revisions of the Manual. In this way, engineers could become involved in Manual production. Following this logic further, it could mean that maintenance engineers could potentially create new entries for the Manual by video-recording a particular sequence of tasks, with appropriate commentary (which could be transcribed to text via speech recognition), and submit this to the MPA.

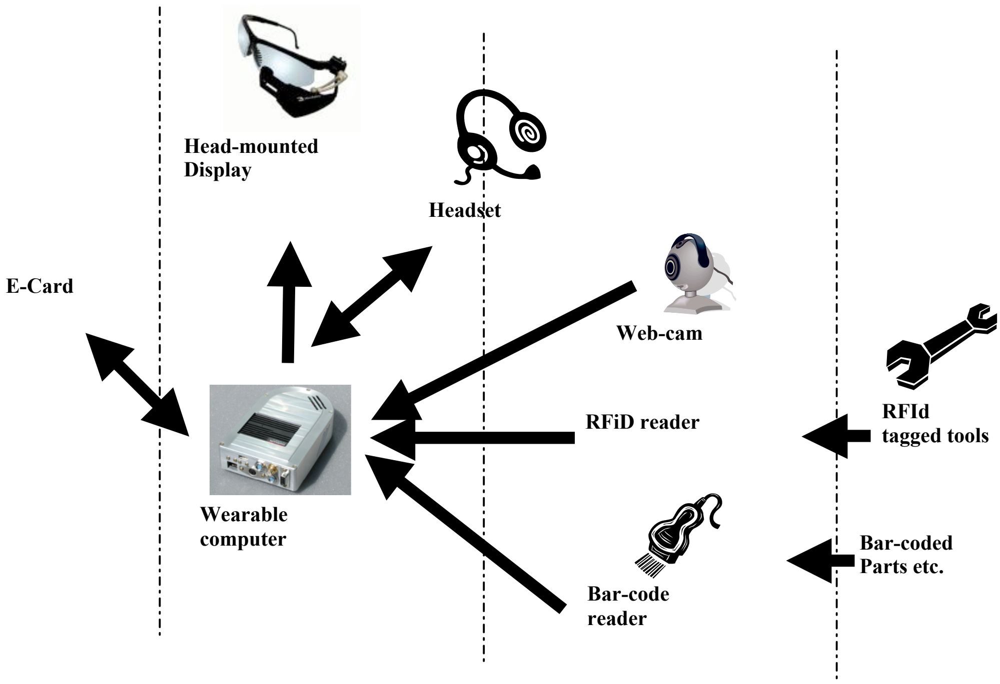


FIGURE 1: schematic of prototype

3. CONCEPTUAL DESIGN

The goals of the work reported in this paper were three-fold: (i.) present E-card on a head-mounted display; (ii.) manage E-card over a wearable computer; (iii.) use actions in the world to control a computer. The first two goals required the ability to request, read and translate XML script so that it could be acted upon by the user and could be read out by a speech synthesiser. While this turned out to be a non-trivial task, it will not be discussed further in this short paper. The third aspect, using actions in the world to control a computer, represents an innovative approach to interacting with wearable computers, particularly in the context of maintenance work. Figure one shows a schematic of the basic architecture of the prototype. The basic concept is that the content of a work card is drawn down from E-card for a particular engineer. Engineers would scan the bar-code on their ID to log on to the system (which is a feature already supported by E-card). This would provide an initial record that Engineer X was using procedure Y at time and date Z. The E-card presents the user with an initial screen with a Summary of the work to do. The user can then ask for a list of Parts and Tools. After checking that all the required tools are to hand, the user works through the task steps.

4. FUNCTIONAL PROTOTYPE

The prototype runs on the x^4 wearable computer, developed at The University of Birmingham. This has a Pentium III processor with 128Mb hard-disk. All external components are interfaced via USB to the computer (in some instances, this requires serial to USB conversion). The application is written in Visual Basic .net 2005 and runs under Windows XP. Figure two shows the system being worn and used during a demonstration.

FIGURE 2: demonstrating the prototype

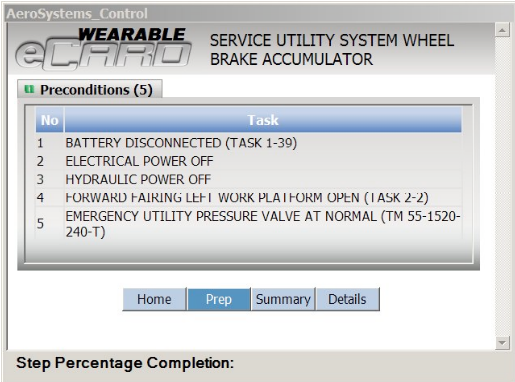


FIGURE 3: preparation page

The Wearable E-card manual presents an introductory page to the user. From here, the user navigates to the Preparation page (figure three) which lists the activity required prior to commencing maintenance work. From here the user can either access a Summary page (which lists the tools, components and people related to the procedure) or to the Details of the procedure. As a page of the manual is accessed, its XML code is stripped down to define tools required, procedural steps to follow, and warning messages. This is then used to define both the information presented to the user and also to monitor activity, e.g., if the user picks up a tool that is not mentioned in that task step, then a warning message is displayed. Thus, the code builds up a hierarchical model on the fly of the steps including the necessary conditions for step completion.

The XML pages are requested along with the HTML pages to display to the user. This is done by constructing an URL as follows:

New_URL = ecard + pipe + "?screen_mode=" + screenmod + "&taskid=" + taskid + "&step_number=" + newstepnum + "&sequence_number=" + seqnum

The software keeps tabs on the step number, sequence number to allow navigation through the work cards. The XML page is parsed by serialising the data and stepping through it in sequence. When a node name is located code runs to parse out the data and store it in a data structure held in memory. The node names for a step are: {Step, Warning, Caution, note, diagram, training, name, number, text, and last_step}.

Code also builds up a list of summary data, such as tools, parts, and people required to accomplish the task. This is held in memory as the task progresses. When objects in the real world are used, the object is compared against each step's data to determine a level of completion. Summary data is held in separate XML files that have nodes: {Summary, people, with descendant person, parts with descendant part, tools with descendant tool, step_transistions with descendants from_step and to_step}.

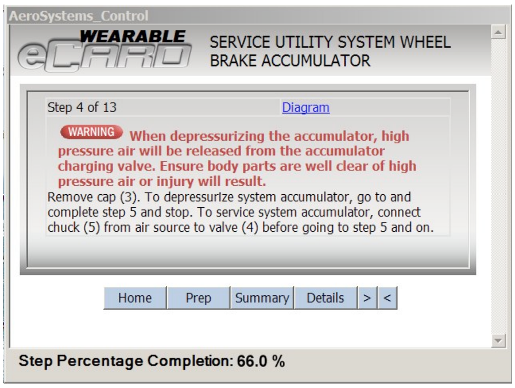


FIGURE 4: Screen shot of application

Figure four shows the task of depressurizing the accumulator. Using the spanner to remove the cap, would lead to 33% of the task being completed and interacting with the chuck would increase this to 66%. The system then requires the user to say "confirm" to complete this step of the procedure, which would then move the system to the next step. On completion of the procedure, a task report is saved to file. This logs all of the task steps viewed, together with components scanned and tools used. In this way, a complete record of the procedure can be automatically captured. This can form the basis of the maintenance log that can be signed off by the engineer.

5. DISCUSSION

In this paper we present the concept and prototype of a context-aware maintenance support system. The main innovation in this work lies in the use of tagged tools to provide input to the computer. In this way, real world activity that involves the use of tools to maintain components (together with recognition of those components through scanning of bar-codes) forms the basis of human-computer interaction. Previous work [6] demonstrated that it was possible to recognise actions (through the use of hidden markov models applied to data collected from accelerometers on the person), and work is currently in progress to link this action recognition process with the object recognition that can be obtained through RFiD. This will allow reporting of not only what tools a person picked up, but also how these tools were used. Ultimately, the aim is to have a means of logging a person's activity implicitly, by monitoring the objects they interact with and the actions they perform. We have been approaching this question as purely a technological challenge, and we believe that the prototype reported in this paper goes a long way answering the challenge. However, the notion of monitoring people (through such ubiquitous computing) raises ethical concerns (not least related to questions of surveillance at work) and also questions of capture, e.g., if a person picks up a tool and then puts it down, should this be logged? Our answer to these concerns relates to the challenges associated with paperless working and zero-error maintenance – if it is possible to write a report simply by doing the job, and also to have some means of ensuring that errors are picked up while the work is being done, then this might make the technology an attractive proposition.

ACKNOWLEDGEMENTS

This work from the Human Factors Integration Defence Technology Centre was part-funded by the Human Sciences Domain of the UK Ministry of Defence Scientific Research Programme.

REFERENCES

[1.] Bass, L., Siewiorek, D., Smailagic, A. and Stivoric, J. (1995) On site wearable computer system, *CHI'95*, New York: ACM, 83-84

[2.] Mizzell, D. (2000) Boeing's wire bundle assembly project, In Barfield, W. and Caudell, T. (eds.) *Fundamentals of Wearable Computers and Augmented Reality,* LEA 447-470

[3.] Nilsson, J., Sokoler, T., Binder, T. And Wetcke, N. (2000) Beyond the control room: mobile devices for spatially distributed interaction on industrial process plants, In P. Thomas and H-W. Gellerson (Eds.) *Handheld and Ubiquitous Computing: 2^{nd} International Symposium, HUC 2000,* Berlin: Springer-Verlag, 30-45

[4.] Casner, S. and Puentes, A., 2003, Computer and broadband technology in the aviaition maintenance workplace, In *FAA General Aviation, Aviaition Maintenance, and Vertical Flight Program Review*, September 10-11, Reno, NV

[5.] Stein, R., Ferroro, S., Hetfield, M., Quinn, A. and Krichever, M. (1998) Development of a commercially successful wearable data collection system, *Proceedings of the 2^{nd} International Symposium on Wearable Computers,* Los Alamitos, CA: IEEE Computer Society, 18-25

[6.] Schwirtz, A. and Baber, C. (2005) Smart tools for smart maintenance, *IEE and MoD HFI DTC Symposium on People and Systems,* London: IEE, 145-153

Integrating Social and Cultural Variances into International eCommerce Interface Design

Jainaba Jagne[1], Serengul Smith[2], Paul Curzon[3], Bob Fields[4]
[1, 2, 4]School of Computing Science, Middlesex University, The Burroughs London NW4 4BT
[3]Department of Computer Science Queen Mary, University of London, Mile End Road, London E1 4NS
[1]j.jagne@mdx.ac.uk [2]s.smith@mdx.ac.uk [3]pc@dcs.qmul.ac.uk [4]b.fields@mdx.ac.uk

Designing software and sites for different cultures requires not only translation of native language and terminology but also translation of metaphors, icons, colours and other semiotics. Mismatches that contradict a user's expectations of metaphors tend to create cognitive conflict, which in turn leads to decreased performance levels. The correct use of metaphors is therefore dependent on understanding the mental model the user has of that concept. This paper addresses the integration of cultural and social factors of metaphors into cross-cultural interface design. It also addresses the emerging need to look into the doctrine behind the metaphors being used. We investigate how people from different cultural backgrounds shop and how their metaphors can be translated into eCommerce interface design. Results obtained show that interface designers not only need to look into heuristics and translating aesthetically related issues but also deeper cultural understandings, perceptions and beliefs of their target audience and market.

Culture, e-commerce, interface design

1. INTRODUCTION

Consumer behaviour is a process with many fundamental influences such as personal stimuli and a variety of social influences [3]. Studies carried out by Russo and Boor [4] and Zahedi, Van Pelt and Song [7] into the impact of new technologies suggests that users show resistance to, and reject products with, Western metaphors in favour of products localised according to their cultural customs, idioms and so on. In the field of eCommerce this is especially significant due to the exponential growth of the Internet. It is now easier for customers to change buying preferences and transfer to a different brand more localised to their cultural background [5].

The study described in this paper aims to provide insight into the doctrine behind the concept of the shopping to eCommerce metaphor. The results obtained will be beneficial not only to eCommerce interface developers but to all developers of interfaces that are designed for multi or cross-cultural use that need a metaphor translated *e.g.* libraries to the digital library and banks to the online banks.

2. METHODOLOGY

Empirical work was carried out in this study using a multi-method approach comprising of interviews and observational ethnographic studies. The use of a multi-method approach added to the rigor of the results obtained and ensured that they were not biased due to any one method being used. It also provided various types of data in the form of notes, audio recordings, etc. The validity of the results was further heightened by procedural triangulation [1].

Interviews for the study were conducted with thirteen subjects from the United States of America, Britain, Ghana, Cote D'Ivoire, Nigeria, Senegal, Turkey, China, India and The Gambia. The subjects were all conducting research in Britain, were computer literate and were of similar demographics in terms of age, financial/social standing and educational level. The interviews were conducted to find out about their shopping characteristics and to understand whether these characteristics were based on or biased by culture. Each subject had experienced shopping in different countries and was therefore able to compare and contrast differences in shopping behaviour between the countries. The categorisation and analysis of the interviews was informed by grounded theory. However, more interviews will have to be conducted to provide a more equal representation amongst the subjects in terms of nationality and gender to add validity to the categories.

Observations were carried out on the indigenous population in Britain and The Gambia. So far the ethnographic studies have only been conducted in the smallest retail outlets. In Britain this was a ‘corner shop’ in Finsbury Park, London and the equivalent, known as a ‘boutique-e-narr’, in Fajara, Banjul.

The emergent themes approach [6] was used on the raw data to form the categories described in the results section. Firstly, the data from the interviews and observations were synthesised. Secondly, the data was then indexed and structured to find broad patterns. Thirdly, new structures and themes were then formulated.

The results described below were also scrutinised and affirmed by a panel of international experts from multiple cultures.

3. RESULTS

First we outline example observations for illustrative purposes before outlining the results emerging from our analysis of the data collected.

It was observed in Finsbury Park that the customers entered the store, browsed for the required product and went to the till to pay. There was hardly any spoken communication between the customer and the shopkeeper other than the confirmation of the total price.

In The Gambia however, the customer did not have access to the goods as they were partitioned. Customers walked into the store and greeted the shopkeeper and people sitting in the vicinity of the shop. They then asked for the required goods and were told the price by the shopkeeper. The money was then handed over and thus ended the transaction. Explanation for the observations on the need to greet in The Gambia is given by Colley [2]:

> The exchange of greetings is the key to successful interaction with Gambians at every level, whether in the market, on the street, in the office, or over the telephone. People are taken aback if you do not greet first before beginning a conversation even if you just want to ask a question. You may also be baffled to see a Gambian doing something you consider “really important” stop everything to spend ten minutes greeting a friend he has seen just hours ago. But once again it is because greeting acknowledges the existence of another human being and taking the time to relate to him or her in a personal way is a priority in Gambian society... [2]

Another noted difference between Britain and The Gambia was the refusal of the shopkeeper to sell certain goods to the customer. In Britain a sign was observed that stated that tobacco and alcoholic items would not be sold to customers below the age of eighteen. This is due to it being the law. In The Gambia one notable piece of conversation was a young lady who asked for a razor blade. The shop keeper refused to serve her the item and when she queried this, she was informed that they do not sell razor blades at that time of the day. Follow-up research into the Gambian culture showed that items like razor blades, needles, soap and charcoal are not sold after twilight hours for superstitious reasons [2]. Further observational studies are to be conducted in Britain, The Gambia and America and at varying locations and store types.

The results of our research show a distinct line of cultural commonalities and differences between the subjects. The categories of shopping behaviour created were termed as culturally independent; characteristics shared across cultural boundaries and culturally dependent; characteristics displayed by certain cultures only.

Culturally independent – The following were commonalities found amongst the subjects, which were irrelevant to the countries they came from and their cultural background.

- **Data security.** All showed concern about having to register on sites and having their data being accessed by companies.
- **Online images versus catalogue images.** Colour contrasts rather than tactile issues were the main reason why fashion items like clothes and shoes were bought from catalogues but not online.
- **Gender specific division of labour.** Men in the study tended to dislike shopping for what they deemed as feminine items such as general household items and clothes, but were eager about shopping for electronic items and cars etc. Whereas women generally did not show interest in those items but were always enthusiastic about shopping for clothes, food etc.

Culturally dependent - These factors were heavily influenced by the subject's national and cultural background.

- **Trust in corporations.** The British and the Americans especially had a deep embedded mistrust of large corporations whereas the other subjects still had some faith in big conglomerates. The quotes below are from the interviewees talking about shopping for groceries from big supermarkets online.
  American – "*What… they don't care. They are… they are just selling a commodity, so () make a lot more sense for them to go through... to just shove whatever they have out the door….. Why would they care? [laughs] No. I don't. Even if they go through and have that as a corporate policy, then I know that they are not going… its not going to be the manager who says, 'We want to go through and treat our customers valuably'. They are gonna pay somebody, two or three pounds, you know... they are gonna pay somebody who doesn't speak any English to go through and just pack up boxes based on numbers or based on pictures. So they can go through and justify the cost. I don't trust corporations.* "
  Gambian – *"I probably would. Probably would… but then you would think that if you have a particular supermarket that's gone through the trouble of doing that… setting that type of thing up, then obviously, whoever is going to be out there choosing your fruits for you should be doing it properly..."*
- **Attitude towards shopping.** Subjects from the less economically developed countries like India, Cote D'Ivoire and The Gambia found shopping to be a rewarding experience whereas the Westerners from America and Britain found it to be another everyday activity and some even regarded it as a chore. When asked about what they liked about shopping, the non-Westerners said it was the whole "experience" of shopping but the Westerners spoke of the particularities of a shop layout or stock variety.
  Gambian – *"So you go in there ... its that whole experience. Its the whole package really. Aside from going and getting what you want, I mean, there is other things... just you know, the experience."*
- **Family involvement.** Without prompting during the interviews, the subjects from India, The Gambia, Senegal, Nigeria and Ghana brought up family issues and said purchase decisions were family oriented the majority of the time. The Westerners from Britain and America were more autonomous when it came to their shopping habits and did not mention any family involvement at all.
- **Economic division.** There is a wider and sharper gap between the better off and the poor in the less economically developed countries such as India, Senegal and The Gambia. This meant that luxury items were only affordable by the more financially well off. In the more economically developed countries like Britain and America, the production and availability of identical items with various degrees of quality and therefore expense, ensures that the less well off were still able to have access to all the items the social elite have access to. This was observed and confirmed during the interviews.

4. DISCUSSION

The results have implications as to how sites are designed on a generic and localised level, such as:

- Adding family oriented aspects to the interface of sites localised to India, The Gambia, Senegal, Nigeria and Ghana. Subjects from these countries have shown that family decisions are of significant importance.
- Scrapping the registration forms from all sites; this is no longer done in America but is still common practice in Britain. All the subjects displayed displeasure at having to provide their personal information online.
- Removal of conglomerate stamps from British and American sites. The loss of trust in major corporations means that customers would be less likely to shop from a site that has a prominent reminder of the corporate brand.
- Conducting further research into colour projection on all sites. Images on visual display units and their correlation to the actual artefact are of vast importance if the success of catalogue ordering is to be matched by online shopping.
- Prominent salutations should be included on the Gambian sites. Greetings are important to Gambian people as a form of validating the existence of customers.
- Further research has to be done into what items are taboo to sell in certain cultures.

This work is ongoing and the results will be translated into a set of comprehensive design guidelines.. The processes and the influences of shopping behaviour can be understood further through the ongoing research of this project. Perfect prediction of what interface will work with each culture is never possible, but continuation of this research and use of its results will aim to mitigate the risks of marketing failure and alienating or insulting customers.

5. CONCLUSION

As mentioned earlier, several studies carried out into the impact of new technologies suggest that users show resistance to, and reject products with, Western metaphors in favour of products localised according to their cultural customs. It is therefore important to take the stance of “prevention is better than cure”. In other words instead of letting market forces sort out the problem, looking into the culture might prevent the company from losing the money first. In the field of eCommerce this is especially significant due to the exponential growth of the Internet. It is now easier for customers to change buying preferences and transfer to a different brand more localised to their cultural background [5] by just clicking a button.

REFERENCES

[1] Denscombe, M. (2003) The good research guide for small-scale social research projects. Glasgow: Open University Press

[2] Colley, E. (2002). Gambian culture notes. [Online]. St Mary’s College of Maryland. Available from: http://www.smcm.edu/academics/hss/studyabroad/gambia/GambiaHB/pdf/12_2002Gambia_culture.pdf. [Accessed 5 July 2005].

[3] Engel, J. F., Blackwell, R. D. & Miniard, P. W. (1993) Consumer behaviour. (7th ed). U.S.A.: The Dryden Press

[4] Russo, P. & Boor, S. (1993) How fluent is your interface? Designing for international users. Human Factors and Computer Systems. Conference (4th. 1993). Proceedings of the conference on Human factors in computing systems. Boston: Addison-Wesley Longman Publishing. 342-347.

[5] Van Dyk, T. & Renaud, K. (2004) Task analysis for e-Commerce and the web in Diaper, D. & N. Stanton, N. A. (Eds). The handbook of task analysis for human-computer interaction. London: Lawrence Erlbaum Associates, 247-262

[6] Wong, W.B.L. (2004). ‘Critical decision method data analysis’. in D. Diaper and N. A. Stanton (eds). The handbook of task analysis for human-computer interaction. London: Lawrence Erlbaum Associates, p.327-345

[7] Zahedii, F., Van Pelt, W. & Song, J. (2001) A conceptual framework for international web design. IEEE Transactions on Professional Communication. 44, (2), 83-103.

Camera Phone Use in Social Context

Hanna Stelmaszewska[1], Bob Fields[1], Ann Blandford[2]
[1] Middlesex University, Interaction Design Centre, The Burroughs, London NW4 4BT
[2] UCL Interaction Centre, University College London,
Remax House, 31-32 Alfred Place, London WC1E 7DP
h.stelmaszewska@mdx.ac.uk, b.fields@mdx.ac.uk, a.blandford@cs.ucl.ac.uk

The primary contribution of this paper is in identifying social uses and practices of camera phone images in co-present settings. Three distinct practices were observed: 'sharing a moment now', 'sharing a moment later', and using photos to initiate social interaction with strangers. We propose that interactive problems with existing systems might be a starting point to generate user requirements for technology supporting social practices in co-present settings.

User experience, HCI, mobile phone technology, camera phone, social interaction, social context, user requirements

1. INTRODUCTION

Personal photography has been a part of the lives of many people for a long time. Photos not only present a documentary of someone's life, but are also of great sentimental value. People use photography to capture feelings, events and personal experiences, and to communicate with others. The latest innovations not only support interactions of people with technology, but also facilitate new forms of interaction with others. Mobile phones and camera phones in particular are examples of such technology. In recent years their primary function has changed from being a medium of verbal or text communication to one that uses pictures to facilitate people's social life. Mobile phones with integrated camera and video features have changed the way people communicate, interact and shape their social activities [1, 2, 3, 4, 5, 6]. With the rapid growth in the camera phone market and continued improvements in technical performance, picture quality, and so on, camera phones might supplant the use of digital cameras, in many everyday settings.

Although there has been some research focussing on the use of camera phones [2,3,5,6], there have been no direct investigations of how camera phones are used to mediate social interaction between co-located users. In this paper, we describe such a study of the collaborative use of camera phones by co-located users.

1.1 Related work

In recent years, there has been substantial interest in digital photography, with a particular focus on how the digital medium facilitates sharing of images [7,8,9]. Studies of sharing digital photographs include web-based systems, mobile applications and multimedia messaging. Most of the studies focus on personal applications for sharing images in a remote location [1,2,9] but work on sharing images in co-present settings is in its infancy.

Kindberg et al. [2] investigated what people capture on mobile phones and what they do with these images. They presented a six-part taxonomy that describes the intentions behind the camera phone images. These encompass affective intentions enriching a mutual experience, communication with absent friends or family and personal reflection or reminiscing; and functional intentions supporting different tasks: mutual, remote and personal.

A field study conducted by Kato [1] explored how use of mobile phones and camera phones changes people's daily activities in Japan. He argues that the new ways of pervasive photo taking enabled by camera phones allows people to document their lives on a daily basis, which can be preserved and shared as a life of local community.

The work of Okabe [5] focused on practices of Japanese camera phone users, which include personal archiving, intimate sharing, peer-to-peer news and sharing. He argues that capturing and sharing visual information cannot be divorced from the social relations and contexts. This is in line with Scifo's [6] research, which shows that taking photographs on camera phones and using MMS communication allow younger users to establish their identity within social groups and can intensify communication within that community.

The relevance of social relations and uses of photographs were also identified by Van House et al. [9]. They discovered five different sets of social uses of personal photos: creating and maintaining social relationships, constructing personal and group memory, self-expression, self-presentation, and functional communication with self and others.

Photos could also be used for social interaction. A mobile picture system (MobShare) developed by Sarvas et al. [10] supports that by transferring photos from the phone to different devices allowing people to share and discuss photos. These include transfers (1) to another phone over the network (e.g. MMS), (2) to a PC, (3) to a network server over the network, and (4) to a printer using a cable connection or Bluetooth.

This study is specifically concerned with people's experiences when using camera phones for social interaction in a co-present setting. The paper draws from an earlier more general study into people's experience and emotions using personal technologies such as PDAs, digital cameras and mobile phones [11].

2. DESCRIPTION OF THE STUDY

Many methods have been used to study people's uses of mobile phones. For instance diaries, interviews and field studies [1,2,3,5,10]. Because we wanted to obtain insights into the ways people use their camera phones as a medium for social interaction, we adopted Kindberg at al.'s methods [2] of asking participants about circumstances and reasons for taking these images and their life cycle. In addition, we conducted a set of field observations to develop a better understanding of people's practices using camera phones.

A group of seven students including undergraduates, PhD students and college students, aged 18–27, was interviewed. The interviews lasted between 25 and 45 minutes and were recorded and later transcribed. Each of the subjects had at least a year's experience of using a camera phone. The participants were asked to describe how and for what reasons they use their camera phones. They were also asked to show a few of the images (pictures or video) stored on their phones and explain the setting in which each image was produced and shared with others. Participants were encouraged to describe the circumstances in which pictures were taken, who took them, whether were they taken by the participants themselves or received from another person, and the reason for taking these pictures. Furthermore, the interviews also enquired into the storage, and transmission of images: how, why, and to whom images were sent.

The data from the field observational studies were gathered from a variety of different public spaces including: pubs, restaurants, leisure and entertainment places, museums, and public transport (tube and buses). There were ten individual instances captured, monitored and recorded in field notes. Data from both studies (interviews and field observational studies) was transcribed and then analysed by firstly, coding the data by using qualitative methods to identify emerging themes; and secondly the themes were merged to extract the high level of concepts that gave the outline of the use and practices of camera phones.

3. FINDINGS

The field observation study revealed many instances of people being engaged in social interaction using camera phones in co-present settings. The in-depth interviews provided extended information to support these phenomena and the following sections will discuss them in detail.

3.1 Social uses of camera phones

Consistent with other studies [2,3] we found that people take photos for individual purposes, which are evocative of special events, trips, holidays, or beautiful landscapes. A general practice is to share them with friends and family, which establish their social interaction [11]. Sharing digital photos is often done remotely via email or by posting them on the web [11,12]. However, we observed three specific practices that occur between participants who are co-located. These are 'sharing a moment now', 'sharing a moment later' and using photos to initiate social interaction with strangers.

3.1.1 'Sharing a moment now'

This study demonstrates that one of the practices of using camera phone was to take a 'spur of the moment' photo or video and share it with people in co-present settings (i.e. present at the scene). People reported having fun when taking photos or videos of their friends' behaviour, and then viewing them collectively in-situ. This kind of behaviour seems to motivate and shape social interaction as one of the participants (A) reported:

> ' … she was happy and funny (referring to a friend) … far too engaged with dancing to notice what was happening around her … and I just thought that I'll just take that picture. … there were few of us friends so then I showed them

and then other friends were taking more pictures of her dancing and we were waiting for her to realise what was going on … we were all taking pictures of her … we shared all the pictures and picked up the funniest ones. It was so funny because she couldn't believe that we did that and she didn't even notice it.'

Data shows that photos were used for functional purposes as well, which is consistent with [2,9]. It was observed that people took a picture of a map displayed by the leader and then viewed it on their camera phones. This allowed every person within the group to see the map clearly and use it for further reference.

Another common practice was to transfer photos across phones using the Bluetooth technology so everybody concerned could store and use them when needed. However, it appeared that some people found it difficult to use it and either abandoned the transfer or asked friends for help.

3.1.2 'Sharing a moment later'

A social interaction at co-present settings was reported to be associated with participants' experience when viewing pictures/videos stored on individuals' phones but taken previously. The intention behind such activity was to share memories of special events, report on events to those who were absent at the time of events, create or share a documentary of a friendship or family life. People were more inclined to use photos for storytelling, which is in line with [2,8]. However, since phone screens are very small (remarked upon by participants) it was common to use other media like computer or TV to display photos in order to improve visibility of pictures and enhance the experience of their viewers.

'I transferred them onto my computer … I'm quite organised with my pictures so I categorise them and put them in kind of albums and sometimes when I'm with friends we like to go through pictures and have fun.'

was reported by the participant A whereas participant M commented:

'…sometimes what we do is we Bluetooth to transfer our pictures to one of our computers and then have a slide show so everybody can see it. …you see the phone screens are very small and if we all want to have fun we need to see those pictures simultaneously. With camera phones we can't see it clearly if there are more then two or three people looking. It's just not enough space …'

Sharing photos at co-present settings proved to be a way of social interaction that brings fun and joy to people's lives. An extract from an interview confirms this point:

'I'll show them (referring to family) what I managed to capture and then we have a good laugh.'

or another comment from the participant A:

'… you take pictures and when you view them you can laugh and have fun'.

3.1.3 Social interaction with strangers

As reported in our previous study [11] social interaction can take different forms from text, graphics, to interactive games. All of them occur between friends or members of family sharing the same technology (i.e. computer, digital camera or mobile/camera phone). The most striking findings were the camera phones being used as a new medium for initiating social interaction with strangers. It was conveyed that people take photos of others (who they like) to show their interest, introduce themselves, or simply start a new social relationship.

The comments from one of the participants' (E) support this claim:

'I was at the Harvester, a restaurant/pub thing, …and there was a small window with glass between it looking like a fake door and the guys were looking through that doing (mimicking facial expressions) and then I saw one holding his camera phone against one of the window things and there was a picture of me going (shows facial expression) and I didn't know that they were taking it … I didn't really mind. It's a good humour… it was kind of friendly, sort of vague flirting without talking … just taking pictures.'

or another remark by participant M:

'We were in the bar … having fun and there was this guy dancing [laughing] kind of a very funny dance … almost like an American Indian kind of dance … and one of the girls from our group took a photo of him because she liked him and she was showing it to us so instead of looking at him we could see his picture … and when he saw her taking pictures of him he did the same to her… the whole situation was funny … at least we had fun watching them two taking pictures of each other instead of talking …'

This kind of behaviour typically occurred in public spaces such as pubs, bars, or clubs where people usually gather for social events and interaction with others is a part of entertainment. In our study, the social interaction took place through digital photos. Though, this was not always appreciated by those involved. Some participants felt uncomfortable and annoyed with those taking photos without permission or agreement, as participant L noted:

> 'I don't know if I would be offended so much. I think it depends what for … sometimes you get photographers going like around pubs and clubs … and I never said yes to the photo. The other night when I was there with my friend and this group of guys we met before errr … this guy said: 'Oh yeah, lets get a picture' but we went like:' no, we really don't want to'. And they had one done anyway and this kind of annoyed me a bit because … it's fair they wanted the picture of us but we didn't really want to be in it. … I think it depends how much choice you are given as whether or not you want a photo taken'.

It seems that communication does not only take place via technology but also alongside it, which is consistent with Stelmaszewska's et al. findings [11]. Moreover, Van House et al. [9] argue that technology (e.g. online photo blogs) is used to create new social relationships. Although this study is at its primary stage and further evidence is required, we suggest that camera phones provide new channels of social interaction within co-present settings.

3.2 Interactional problems

In general, the data illustrates that although camera phones are becoming a part of our social lives there are interactional problems that hinder people's experience. When using camera phones for 'sharing a moment now' activities sending photos from phone to phone is a common practice. However, several participants encountered difficulties when using MMS feature. In addition, using Bluetooth facilities involved a lot of settings and it was not easy to find the function on a phone.

Another important issue revolves around the lack of compatibility between camera phones. People do not send pictures to their friends because they will not be able to view them was commented by the participant E:

> '… none of my friends are really doing this … you have to have the same phone or something to be able to send it and for them not to just say: 'message not being able to deliver' or whatever. Some people tried to send pictures on my phone but I never got them.'

Another highly valued property for the majority of participants was the quality of pictures. They claimed that with better quality pictures they could print and exhibit them in their living environment instead of storing photos on their phones or transferring them on a computer, or the Internet.

A quick access to photo image features was a further crucial issue pointed out by participants; one remarked:

> '… one of my friends helped me to set it up so I can us it by pressing just a couple of buttons instead of going through menus and stuff. It was horrible. I missed so many great pictures because of that and I was very upset about it. … it's very important. I could have so many great pictures but couldn't find the camera function on my phone … it was very frustrating.'

Finding archived photos that is a predominant activity in the context of 'sharing a moment later' appeared to be another concern. As observed during the field studies, people do not want to spend too much time looking for pictures when engage in social interaction. This caused frustration and dissatisfaction as said by participant J:

> 'Where is it?!!! S… Hrrrrrrrrrr.'

So providing functionality that is transparent to users might reduce the number of interactional problems, which is of paramount importance when designing systems. It might also enhance the use of camera phones creating pleasurable experience. This could be a part of user requirements that clearly indicates what is required from camera phones design to satisfy not only functional purposes but also generate experience that would evoke pleasure, joy and fun mentioned by participants as an essence of social interaction.

4. DISCUSSION AND CONCLUSIONS

This paper has described distinctive practices of camera phones users occurring in co-present settings. We have argued that camera phones enable new ways of social interaction through taking and sharing photo imaging and these activities are inseparable from social relations and context, which is consistent with Okabe's [5] and Scifo's [6] work. However, system designs for camera photos need to overcome challenges of photo transfers, quality of photos, and quick and easy access to required functions, to name just a few.

More generally, when designing camera phones that are used for social interaction it is crucial to develop an understanding of emerging uses, practices and social activities that are supported by camera phones. Moreover, identifying interactional problems within existing systems might be a good staring point for discussing user requirements allowing designers to develop systems that would fulfil utilitarian as well as user experience needs.

REFERENCES

[1] Kato, F. (2005). *Seeing the "Seeing" of Others: Conducting a Field Study with Mobile Phones/Mobile Cameras.* Paper presented at the T-mobile conference, Budapest, Hungary, 28-30 April.
[2] Kindberg, T., Spasojevic, M., Fleck, R., & Sellen, A. (2005). An in-depth study of camera phone use. *Pervasive Computing, Vol.* ***4*** (No. 2), 42-50.
[3] Kindberg, T., Spasojevic, M., Fleck, R., & Sellen, A. (2005). *I Saw This and Thought of You: Some Social Uses of Camera Phones.* Paper presented at the CHI'05, Portland, Oregon, USA, 2-7 April.
[4] Ito, M., Okabe, D. (2005). Technosocial Situations: Emergent Structurings of Mobile Email Use. In Ito, M., Okabe, D. & Matsuda, M. (eds.), *Personal, Portable, Pedestrian: Mobile Phones in Japanese Life.*
[5] Okabe, D. (2004). *Emergent Social Practices, Situations and Relations through everyday Camera Phone Use.* Paper presented at the International Conference on Mobile Communication, Seoul, Korea.
[6] Scifo, B. (2004). *The Domestication of Camera-Phone and MMS Communication: The Early Experience of Young Italians.* Paper presented at the T-Mobile Conference, Hungary.
[7] Frohlich, D., Kuchinsky, A., Pering, C., Don, A., & Ariss, S. (2002) *Requirements for Photoware.* Paper presented at the CSCW'02, New Orleans, Louisiana, USA, 16-20 November.
[8] Balanovic, M., Chu, L., & Wolff, G. J. (2000). *Storytelling with digital photographs.* Paper presented at the SIGCHI 2000.
[9] Van House, N., Davis, M., Ames, M., Finn, M., & Viswanathan, V. (2005), *The Uses of Personal Networked Digital Imaging: An Empirical Study of Cameraphone Photos and Sharing.* Paper presented at the CHI'05, Portland, Oregon, USA, 2-7 April.
[10] Sarvas, R., Oulasvirta, A., & Jacucci, G. (2005). *Building Social Discourse Around Mobile Photos - A Systematic Perspective.* Paper presented at the MobileHCI'05, Salzburg, Austria, 18-22 September.
[11] Stelmaszewska H., Fields. B., & Blandford, A. (2005). *Emotion and technology: an empirical study.* Paper presented at the Emotions in HCI design - Workshop at HCI'05. Edinburgh
[12] Counts, S., & Fellheimer, E. (2004). *Supporting social presence through lightweight photo sharing on and off the desctop.* Paper presented at the CHI, Vienna, Austria, 24-29 April.

Older People's Experiences Route-planning

Judy Wilson[1] and Paul Curzon[2]
[1] Middlesex University, The Burroughs, London NW4 4BT, UK
[2] Queen Mary, University of London, Mile End Road, London E1 4NS, UK
j.wilson@mdx.ac.uk, pc@dcs.qmul.ac.uk

Current navigation systems provide clear guidance for moving from one place to another while taking little explicit account of traditional strategies for route planning. This study observed twelve older people working in pairs planning a route with the aim of extracting the benefits to older people for using their life-long procedures. Conclusions include that personal requirements for routes are important. There is also a wide value in cognitive mapping capability beyond immediate navigation tasks. Finally previous personal experience, preferences, memories, geographical and historical knowledge and interest improves the route-planning itself which in turn supports internal cognitive mapping facilities. This research highlights a need for systems to more naturally and easily reflect older users' traditional ways of working.

Navigation systems, fit older users, route-planning, cognitive mapping

1. INTRODUCTION

Navigation is a ubiquitous skill developed through childhood [1] and forms part of the active process of cognitive mapping. It involves the use of both perceived salient features in the environment and secondary sources to aid the ordering, development and use of our internal spatial resources [2] that are frequently supported by external artefacts [3]. Navigation technology increasingly fulfils commonly regarded usability criteria as described by [4]. Systems generate routes in an efficient, timely fashion and allow for some user preference. However, previous work [5] has suggested there are still limitations. Some systems attempt to support personalisation of a route but these require substantial user expertise and persistence.

Our qualitative, user-focused study set out to extract user preferences and benefits when devising and executing routes using traditional, self-selected sources. We consider whether current navigational technology is successfully addressing any such benefits or if their design over-rides long-established values. Our participants can be described as 'fit older users', with a life-time's experience developing their cognitive mapping procedures and processes and a need to remain independent, confident and mobile. Research and design culture not only contain embedded assumptions from which many older people are alienated [4] but also may not fit their long held mental models and internal representations. We consider whether the revealed processes give 'added benefit' not embedded in the envisioned contextual use of navigation systems and make suggestions for requirements for future design.

2. METHODOLOGY

We conducted a qualitative, un-timed scenario based study. Twelve people aged 53-88 were videoed, using artefacts of their choice planning a route from their local vicinity to a restaurant that required them to cross the M25 boundary into or out of London. In all instances a pertinent 'A' road was 'banned due to road works' and the routes involved at least one borough of London and a county outside. Six participants were also observed carrying out the journey. Unpaid pairs of self-selecting friends who were not normal travel partners were used to reduce anxiety and encourage discussion. Route planning is often a social activity involving conversation and assistance. Tversky *et al* [6] stresses the role of language in the development of spatial mental representations and we aimed to encourage verbal exchanges in a relaxed 'everyday' setting. Lave [7] highlights the effect of context on cognitive phenomena and the importance of natural settings so the planning was in the home of one of each pair. Role-play activities enable a 'closer to event accounting' [8] than asking participants to recount and replicate passed events. The activities also triggered the recall of previous events and pertinent associations and justifications for their behaviour and these were explored during post-activity interviews. They were not normal travel partners adding a level of distortion.

Identifier	Sex	Age	Driver	Usual role in car	Plans routes	Undertook planned route
Pair A P1	F	83	No (but until 1 year ago)	Passenger	Never	As passenger
Pair A P2	F	53	Yes	Driver/Passenger	Yes	As driver
Pair B P3	F	83	No (drove until recently)	Passenger	Never	
Pair B P4	F	71	Yes	Passenger/Driver	Yes	
Pair C P5	M	66	Yes	Driver	Yes daily	
Pair C P6	F	66	Yes	Driver/Passenger	Yes	
Pair D P7	F	88	Yes	Driver	Yes	As passenger
Pair D P8	F	72	Yes	Driver/Passenger	Yes	As driver
Pair E P9	M	62	Yes	Driver/Passenger	Yes	
Pair E P10	F	72	Yes	Driver/Passenger	Yes	
Pair F P11	F	68	Yes	Driver/Passenger	Yes	As passenger
Pair F P12	F	60	Yes	Driver	Yes	As driver

TABLE 1: Participants of Study

3. RESULTS

3.1 What matters in routes is personal

Participants revealed a range of highly personal requirements for routes reflecting their own desires and needs: aesthetics, perceived safety factors (hard junctions, motorways, winding roads), speed, directness, novelty/sameness, levels of congestion and speed camera awareness. Frequently, the desire for speed was over-ruled by other needs. Several participants talked about the view from the car and were often keen to provide themselves with a scenic route.

> "Yes, we've devised what I call a scenic route to get to Bristol…through Thames. Beautiful run but not quick" (P4)

A personal knowledge of particular junctions allows them to select details they wish to avoid.

> "Oh, I used to avoid that Mill Hill roundabout do you know that?…The traffic sort of comes down in three lanes and there used to be nothing to stop it at one stage you just used to sweep onto the roundabouts" (P6)

This is not simply a binary selection between, say, the most direct route and one that avoids motorways, a choice supported by current navigation systems. Rather their routes provided a mix of features that were important on an individual basis. For some, assistance from others may therefore be detrimental if suited to the helper's requirements:

> "I might have done [spoken to someone else] but what I usually find is I live in a family of men and they would all tell me how to do it the way they do it which is quite different to they way I do it" (P6).

Note that other participants, for a variety of reason, clearly did like discussing routes with others.

3.2 Is cognitive mapping important?

Internal cognitive maps let us organise our spatial understanding and knowledge and allow us to 'focus our attention on different categories of information' [3]. These internal representations include where something is and also what occurs there and when. The level of importance given to categories of information reflects our individual choices and interests at a point in time but is also reliant on an internal representation that we can reorder as required. These internal resources may rely on varying external resources and features to be useful in some or all contexts.

Many of our participants displayed an intrinsic pride in having a good cognitive map and their ability to think spatially.

> "…I really do know London pretty thoroughly and not just the destinations and the Elephant and Castles and the Hyde Park Corners, I do know several routes through Mayfair, or anywhere you like to name frankly within that Marylebone Road, river Thames Area….and I can run it my mind like a video again" (P9).

Cognitive map facilities allowed participants to usefully apply their own strategies to planning. At the extreme end of efficiency almost immediate solutions were perfected by re-running parts of the route deemed to need refinement.

> "I went to the destination and in my mind's eye, so to speak, I worked backwards to where we are starting from, to see the most direct route. In a visual sense on a sort of map in my mind's eye and then try to apply the roads that I know to that route" (P9).

Others revealed an attachment to old routes. P10 adapted a route they used 20 years ago partly because it met a desire for a "pretty route". Some had a strong ability to blend internal knowledge with external maps allowing the application of a straight-line strategy (or several straight-lines) then building up detail to fulfil their requirements for a route. Although not expert, these participants were confident and enjoyed exercising their problem-solving capacities. They were blending the newly acquired knowledge with their old internal representations.

> "There's Theobalds, It's north of Theobalds. It's not on this map" (P7).

Historically, Theobalds college was a key local feature. Similarly, when map reading whilst lost en route,

> "Oh my golly, what was Nanny along, Enfield Highway, right that should help" (P2).

Thus incomplete knowledge in the head provided, at a fundamental level, cognitive hooks into externally provided artefacts. Such links to prior experience fed confidence that they knew where they were, what they were doing and could recover if lost. We can categorise these as personalised landmarks or salient elements in cognitive maps.

Internal representations, however incomplete, allow people to decide on roads and junctions that they wish to use and avoid. This has positive outcomes not only in terms of effective route planning but also in terms of their confidence when driving by allowing them to anticipate:

> "I mean if you know a route it makes it that much quicker" (P12) "You can drive nice and fast" (P11). "You anticipate bends and things" (P12).

Participants who believe they know where they are going are confident they will not get lost, or if they do become lost will recover. There was also a genuine satisfaction in knowing where they were going. An alternative approach was displayed by P2 who relied on her ability to respond to a situation as it occurred. Whilst planning her route:

> 'Yes, oh my gosh this is so frustrating [flicking forwards and back 2 pages of the map]. To be honest ...I'd be thinking well, hell I know how to get to Edmonton 'cos I'm just going to follow the back route through Cheshunt, Enfield, Waltham Cross then is Edmonton beyond that?" (P2).

Here however, her difficulties with direction did prove problematic en route but her confidence allowed her to recover (turn round) and deal with the situation (by driving on and referring to a map). Most participants preferred to have a solution 'in the head' before they started, frequently backed up by external elements such as brief notes.

Participants who found difficulties in locating elements in a useful fashion for route-planning still had a collage like representation (see [9]). These multi-viewed images of places with personal meaning and connection revealed a cognitive map that was helpful for deciding where they might go and where their personal stories were located, rather than for deciding how feasible the journey might be or how to devise the route. Most participants revealed a persistent desire to reminisce about past routes and maintain and update their understanding of the world:

> "You see this used to be the main road but now they've made this the main road but it used to be the other way round' (P11).

Passing a new sign for a hotel off the road:

> "As we drive past there's something on the left hand side. Can anybody lean forward because I never have time to read it all." (P11).

The updating of their personal comprehension of the wider world is pleasurable and clearly benefited the route-planning process both before and en route execution. For some participants this was explicit. P5 describes his first experience of travelling by coach on the new M6 toll road,

> "…the first thing I observed and obviously registered up here is that when you leave the M6 old route that goes off to the left and carries straight on to the toll road it's not toll road for the first mile and a half because it joins a junction with the M42 northbound" (P5).

This knowledge is subsequently used when he is driving in the area and road signs would lead him to believe that this bit of the M42 is the M6 and chargeable for toll purposes but he knows this is not so, and uses it without charge.

3.3 Value in the planning itself

By undertaking the route-planning themselves participants are clearly able to ensure a fit with their own personal values for a successful route.

> "It's a reasonable road" (P11). "Yes, its fairly direct" (P12). "Not busy and fairly direct" (P11)

Planning helps to reinforce and build on internal cognitive representations. This reinforcement can provide short-term benefits allowing for immediate serendipitous discoveries of places we wish to locate both within and beyond the scope of the current task. For the longer term an ever-broadening cognitive representation was naturally augmented through the route planning activity,

> "I can picture the golf course now, I couldn't remember" (P12).

Most participants enjoyed route-planning and associated map-reading.

> "I mean I've a very good spatial awareness of maps and things so I think that comes naturally. I mean I can sit and look at maps just for fun" (P12).

Some map-reading involves a general deliberation of an area before route-planning, looking for notable geographical and historical features that may be worth visiting.

Overall, participants' anecdotes and monitored activity suggest that by devising their own route they instinctively provide their own external representation as necessary. This augments memory in a way that meets an individual's needs without the burden of excessive details. It frees them to concentrate on driving with a minimum of distraction.

For the very efficient, years of knowledge acquisition provides an almost video-like reproduction of routes in the head giving a very efficient and timely way of providing, testing and checking a route independent of external resources,

> "I can see almost like a TV camera, I have a camera in my mind and it's running not in fast forward although I will do it in fast forward quite often" (P9).

In general P9 describes their situation as,

> "If I am doing guesswork it's based on an intuition, I sort of know roughly, it's not a total shot in the dark guess. It's a guess based on assembled information in the brain" (P9).

When less certain, this participant would combine their internal representation of the route with the paper-based system similarly described by others using notes or pointing stickers at pertinent points on the map, e.g.

> "If I'm on my own I tend to write things in thick pen and have it on the seat beside me" (P8).

Participants who lacked confidence in their route-planning abilities before or en route still retained pride in past successes and a linkage of personalised features to the current activity. They showed a clear pleasure in talking to family and friends and pride in the abilities of those who assisted them. They tended to drive only a few regular routes or to have given up driving recently. This appeared to reflect a lack of confidence rather than driving ability.

There was some evidence that interactive dialogue (e.g., during a telephone enquiry), provided discussed elements with a dominating salience in participants' internal representations.

4. CONCLUSION

Satellite navigation systems have clear potential to overcome problems for older people, helping avoid social isolation by providing timely, accurate route-information. Whilst none of our participants used such systems they were keen to avoid being lost and remain socially mobile. For people with cognitive impairments, technology can be harnessed to reduce cognitive workload although adapting to new systems may introduce a large initial cognitive overhead. Whilst notable cognitive decline is not apparent for many until deep into old age, Rabbitt [10] concludes that across the population the more fluid capabilities to learn new skills and procedures degenerate more quickly than crystallised cognitive abilities and procedures that are frequently retained deep into old age. Participants were aware of this:

"So what you lack in speed of memory you make up with experience" (P7).

Our study suggests that completely automating the navigation and route-planning task offers advantages but loses benefits of using existing strategies. Whilst technology provides a fast solution to navigating from one place to another this study indicates there are further potential requirements for technology supporting the provision of routes. In particular, there are benefits from incorporating personalised and favoured landmarks and routes as well as acknowledging attachments to known routes. Systems that can learn from common behaviours would be useful to achieve this. Giving ways that participants can interactively contribute to the planning process would not only potentially improve the routes from a personal point of view. It would also help extend and maintain the person's cognitive map. Having a good cognitive map has much wider value than just for navigating from A to B. Providing dialogues in which users can, for example, select the format and level of detail at various points for en route support would also be useful. Pleasures from map-reading could be further enhanced by highlighting and providing links from geographical and historical features. Systems could enable people to remain mobile whilst at the same time allowing them to exhibit detailed personalised control over their route choices; to enable use of internal cognitive mapping processes; and to support user involvement with route-planning. The challenge for our ongoing work is to find ways to do this using natural forms of interaction.

REFERENCES

[1] Piaget, J., Inhelder, B., (1956) *The child's Conception of Space.* Routledge & Kegan Paul, London.
[2] Kitchin, R. and Blades, M. (2002) *The Cognition of Geographic Space.* I.B. Tauris & Co Ltd, London.
[3] Downs, R.M. and Stea, M. (1977) *Maps in minds: Reflections on Cognitive mapping.* Harper & Row, NY.
[4] Jordan, P.W. (1998) *An introduction to usability.* Taylor & Francis, London.
[5] Curzon, P., Blandford, A., Butterworth, R., and Bhogal, R. (2002) Interaction Design Issues for Car Navigation Systems. Proceedings of *HCI: Memorable Yet Invisible, Volume* ***2****,* Sharp, H., *et al* (Eds.) London, August, 38-41 BCS.
[6] Tversky, B., Franklin, N., Taylor, H.A. and Bryant, D.J. (1994) Spatial mental models from descriptions. *Journal of the American Society for Information Science,* **45**(9), 656-668
[7] Lave, J. (1988) *Cognition in Practice.* CUP, Boston, MA.
[8] Yardley, K. (1995) Role play. In Smith, J., Harre, R., Van Langenhove, L. (Eds). *Rethinking methods in Psychology.* Sage pub, London.
[9] Tversky, B. (1993) Cognitive maps, cognitive collages, and spatial mental models. In Frank, A.U. and Campari, I. (eds) *Spatial Information Theory: A theoretical Basis for GIS*, Springer-Verlag, Berlin.
[10] Rabbitt, P. (1999) When age is in the wit is out? In Della Sala, S. (Ed) *Mind Myths: exploring popular assumptions about the mind and brain,* Wiley, Chichester.

Applying Usability Principles to Content for Diverse Audiences

Caroline Jarrett
Effortmark Ltd
11 Bridge Street
Leighton Buzzard
LU7 3AB UK
caroline.jarrett@effortmark.co.uk

Whitney Quesenbery
Whitney Interactive Design, LLC
78 Washington Avenue
High Bridge,
New Jersey 08829 USA
whitneyq@wqusability.com

Ian Roddis
The Open University
Walton Hall
Milton Keynes
MK7 6AA UK
I.Roddis@open.ac.uk

In the summer of 2005, the web team at The Open University faced the challenge of investigating and meeting the needs of a new audience: teenagers leaving secondary schools. It also wanted to reach out to harder-to-find parts of its traditional audience. This paper is about what we learned when re-writing content for these two new groups of users. Our investigation showed that we needed to avoid long blocks of text, to create informative headings, and to provide navigation options within the content. We applied those principles and mostly it worked, but we were surprised by how even minor variations in the language could impair users' understanding. We also found that contrary to some popular guidelines, in-page links and long pages worked well for our users.

Writing, content, users' language, teenagers, web

1. INTRODUCTION AND BACKGROUND

The Open University is a UK-based university that teaches by distance learning. When it was established in 1975, the university was aimed at people knew about university-level education but who missed out it in their late teens/early twenties. Over the years, however, the target market for an OU education has changed. For example, it now has postgraduate and professional programmes, including a well-regarded MBA.

With the introduction of increased tuition fees in the UK, The Open University is now attractive to a new market: teenagers who are planning to go to university straight from school and who might be attracted by the flexibility, opportunity to work while studying, and lower costs of distance learning. This is an audience with whom the OU had little experience.

Meanwhile, the university also asked us to broaden its traditional audience by reaching out to people who had no previous knowledge of, or contact with, higher education. This second audience posed other challenges. It includes both new immigrants and those in the UK without any prior family history of attending university. They are not just choosing a place to study, but learning about higher education and how it works. These "new enquirers" needed support for both the educational and social aspects of their choice.

We started a three-part process:

- Undertaking exploratory usability tests with these audiences
- Searching the literature for relevant research and guidelines
- Developing and testing a prototype web site

Exploratory usability testing

Our exploratory testing showed that we did indeed have some work to do.

Although the main Open University web site contained plenty of information that was relevant to and interesting for our teenage participants, they found it hard to find the specific information they needed. An existing specialist web site that was aimed at teenagers worked well within its remit of providing an introduction, but when it transferred them over to the main site for more information they sometimes became lost or confused. And we found that their tendency to skim-read meant that any lengthy paragraph could throw them. For example, reading a longish page on university fees, they came up with a drastically incorrect price for what the OU costs.

The new enquirers struggled to understand the basic concepts and the vocabulary of higher education. We were not surprised when they stumbled over a technical word like 'pedagogy', but even terms like 'undergraduate' or "B.Sc." were unfamiliar to them and therefore confusing. As they tried to navigate through the course offerings, they were often tripped up by misunderstandings. One particularly telling example was when a participant looking for a course to help her improve her English-language skills ended up choosing an advanced linguistics course called “English Language and Literacy”.

Literature review
Our next step was to find out what the literature told us about how to write and design for teenagers as well as for people so unfamiliar with a complex domain that it posed significant reading difficulties. We also wanted to know whether we should create a dedicated site for each audience, or whether their needs might be met on a single site. In this investigation, we also considered the needs of older adults, because a large number of OU students are over 55.

Quesenbery, 2006 [1] reported on our findings:

> “It’s hard to imagine three audiences that seem further apart than teens, older adults, and people with difficulty reading. We had feared that we would find conflicting advice. But the advice we found was remarkably consistent, even looking closely at specific, diverse audiences. The reasons were different, but the same advice served each of these users in different ways:
> - Avoid long, dense blocks of text.
> - Create informative headings.
> - Provide navigation options within the content”

Developing and testing a prototype web site
Based on this research, we decided to prototype a single site to meet the needs of both “new enquirers” and teenagers. We knew that the information they needed was already available on the current Open University web site, so we undertook an audit to identify appropriate pages. We located 530 pages of possible content to include, some overlapping or repetitive. We designed the information architecture based on this content. Then we re-wrote the content in line with these three principles, and tested it. So did it work?

2. WHAT WE FOUND – ‘AVOID LONG, DENSE BLOCKS OF TEXT’

We knew from our previous research that our audience found many of The Open University’s pages were too complex and dense. We thought of Krug’s third law of usability [2]: “Get rid of half of what’s on each page, then get rid of half of what’s left”. So first, we cut the number of individual pages drastically. Applying Krug’s rule to our 530 source pages, we would have ended up with (530/2)/2 = 132 pages. In fact, at the time of writing the proposed site has about 60 pages.

We found that we needed to be equally determined when cutting the content. For one of our target pages on the site, we started with 35 pages. After choosing the most appropriate bits for our new audiences: around 1300 words. After more editing, the page as tested in the prototype site: 450 words. Not quite Krug’s rule, but close.

But did we lose by cutting? No. The participants said “that’s everything we need”.

Does that mean the original pages were unnecessary? Possibly, but probably not. There is a place for an introduction, like in our new site, and a place for detail. What we wanted was to provide enough information to engage our users and then to entice them to explore the rest of the OU’s web presence in an informed way.

And do the numbers support our contention that our cutting was beneficial? Yes. In our initial study, only one out of 22 teenage participants was able to give an accurate account of what the Open University was and how it worked – and she happened to have a sister who was currently studying at the University. After using the prototype site, all 18 participants were able to do so – even the 15 who had had no prior contact with the University.

Lesson learned: drastic cutting really does work.

3. WHAT WE FOUND – ‘CREATE INFORMATIVE HEADINGS’

When we wrote the prototype, we were rather proud of our initial headings. They were short, used familiar words, and we considered that they should inform our users about what they would find in the text underneath. For example, one part of the site contains basic facts about the OU and explains some of its special terminology. The headings we used for the sections are in the left-hand column of Table 1.

But we found that they were not nearly informative *enough*. For example, we knew that participants would find it much easier to understand the OU if they knew the special OU terms “Courses”, “Points” and “Levels”. So we had a heading: “Courses, Points and Levels”. But participants did not know that they needed to know about the special terms. They skipped over the crucial page to concentrate on ones that were a better match for their goals, such as “Paying for your study”.

This finding was so consistent that after 12 participants, we gave in. We restructured the section to ensure that its headings were both informative and a good match to users’ goals, and incorporated the problematic terms page into the other pages. In our final day of testing, our 6 participants at last did what we expected: visited each of the pages in the order we presented them, no skipping.

Lesson learned: ‘informative’ must mean something that the user wants to be informed about.

Page headings for one section of the web site as tested - first round of prototype	**Page headings for the same section of the web site as tested - final version of prototype**
What is the OU? Courses, Points, Levels Types of courses Types of qualifications What is distance learning? Where is the OU? What you do in a course? When do you study? If you are disabled Entry requirements Paying for your study	What is the OU? Where is the OU? How does distance learning work at OU? What you do in a course? When (and where) do you study? What qualifications does OU offer? How do you apply for entry? What does it cost? If you are disabled

TABLE 1: Headings in two versions of the prototype

4. WHAT WE FOUND – ‘PROVIDE NAVIGATION OPTIONS WITHIN THE CONTENT’

We wanted to "provide navigation options within the content" – but which ones? And what wording?

The current Open University web site makes extensive use of ‘within-page’ (anchor) links, and we had observed participants in usability tests using them without difficulty. So we included them in our prototype. Just as we went to test, Jakob Nielsen published an Alertbox condemning their use [3]. Were we going to be in trouble?

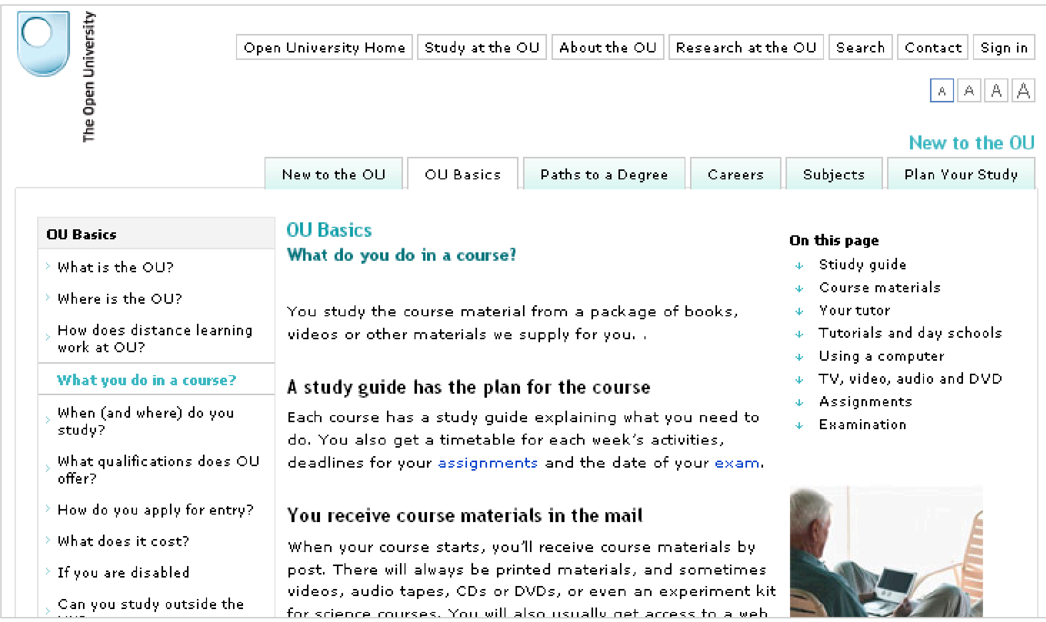


FIGURE 1: Within-page links like these worked well for our participants

We are pleased to report that we were not. Participants used our small, right-hand menus of in-page links, as shown in figure 1, exactly as we expected: they scanned them to gauge what was in the page, and they used them to jump around the page if that met their needs. They swapped easily between using these within-page links and scrolling the page.

We also knew that headings are important as a navigational element, helping readers find information faster. Ginny Redish's [4] article on headings on PlainLanguage.gov suggests that headings must make a connection to the user, either by posting a question, or suggesting an answer. We made each of the page titles a question. Within the page, each section heading is a short statement of one part of the answer to the question posed in the title; the text below the heading elaborates and explains. Because our headings are long, the "On this page" links are shorter, providing a more telegraphic view of the content of the page.

Lesson learned: within-page links and headings can work together to help users learn what is available on each page.

5. WHAT WE FOUND: THE GUIDELINES INTERACT

We also found that it was necessary to consider all three guidelines at the same time: cutting words, using informative words, and getting the right balance of content and navigation. For example, one short paragraph on the home page went through three iterations in three days, as described in table 2, as we struggled to entice participants into one section of the site that we knew would answer many of their questions.

"Clip" from prototype home page	What happened in the testing
1 **Learn the basic facts about the OU** How can a university have 180,000 students and no campus? Learn what it's like to study according to a timetable that suits you.	First version Not bad, but: - No match with the users' top questions - Few participants chose the link Participant "talk aloud" comments showed that they did not understand that the answers to their top questions (what it costs, how it works and what degrees you can earn) would be answered here.
1 **Learn the basic facts about the OU** How can a university have 180,000 students and no campus? Learn how OU study works, types of qualifications offers, what it costs and what it's like to study according to a timetable that suits you.	Second version Better. Many participants want to know 'what it costs' and the words are now there on the home page. But: - The crucial phrases are there but hidden below words that participants did not have in their heads - Few participants chose the link as a first choice
1 Learn how OU study works, what it costs and what it's like to study according to a timetable that suits you. **Learn the basic facts about the OU**	Final version Much, much better. - Participants wanted to know how the OU works and what it costs – and those phrases are prominent on the page - The introductory text now comes before the navigation - Participants chose this link easily: either straight away, or returning to it when they had explored the other sections

TABLE 2: Examples of navigation within content – and what happened to them

6. SUMMARY AND RECOMMENDATIONS

We were challenged to find ways to meet the needs of two apparently diverse audiences: teenagers, and people who had no previous contact with or knowledge of higher education. Our investigation of the literature led us to hope that we could develop one site that would meet the needs of two different audiences, so we created a prototype and tested it.

We found that it did work – provided that:

- we really did follow the guidelines, and continued to follow them through the site iterations
- we listed very carefully to users' needs and how them expressed them, and
- we made sure that the site closely matched both their needs and the language they used to express them.

REFERENCES

[1] Quesenbery, W. (2006) More alike than we think. UX Matters March 2006 http://www.uxmatters.com/MT/archives/000080.php

[2] Krug, S (2005) Don't Make Me Think: A Common Sense Approach to Web Usability (2nd Edition), New Riders Press Berkeley CA.

[3] Nielsen, J (2006) Avoid Within-Page Links http://www.useit.com/alertbox/within_page_links.html

[4] Redish, J. (undated) Headings, Plain Language.gov http://www.plainlanguage.gov/howto/guidelines/headings.cfm

Tools for Safe Colour Selection

Luke Jefferson and Richard Harvey
School of Computing Sciences, University of East Anglia, Norwich, Norfolk, UK.
{laj,rwh}@cmp.uea.ac.uk

A recent Disability Rights Commission Study reveals poor colour accessibility to be one of the most frequently encountered problems by disabled web users, despite numerous guidelines pertaining to colour use on the Web. In light of this, we demonstrate how a simple extension to a standard colour picker can assist with the selection of accessible colour.

Accessibility, colour selection, colour vision deficiency, web content and accessibility guidelines

1. INTRODUCTION

When Henri Mattise wrote of his own use of colour, *"It is not quantity which counts, but choice and organisation"*, he could have been writing about the Web. Over half a century later, despite numerous accessibility initiatives [1], guidelines [2,3], and an extensive usability study that revealed colour to be a recurrent barrier for disabled web users [4], this simple observation is still widely ignored.

Colour is a subtle variable that, when used correctly, can significantly enhance the user's web experience. The careful selection of colour can improve readability, provide structure and increase task performance.

Many designers are keen to use colour, but often the results are counterproductive, and the dangers of misuse are high. Even if designers are compelled to select accessible colours, they may lack the necessary knowledge and experience; colour selection is a complex task that must satisfy aesthetic and practical objectives. The careful treatment of colour accessibility has been a sadly neglected component of web design. For this reason, a critical advance towards addressing colour accessibility would prevent web authors, some of whom are only casual users of colour, from making poor colour design decisions.

Some advanced colour selection tools are available for specific tasks, such as automatic colour palette creation, but few facilitate the creation of colour palettes that are accessible. Those that do, which are mostly only available online [5], simply provide post-evaluation validation. *"A common response from programmers after receiving an accessibility evaluation is that it would have been much easier to incorporate the requested changes at the beginning of the site development lifecycle…building accessibility into a site early in the development lifecycle saves time and money compared to retrofitting."* [6]. This paper demonstrates that by combining a few simple techniques it is possible to avoid these problems and select colours that all users can interpret. Developers can accomplish this with little overhead and be confident that the selected colours are accessible to all.

2. BACKGROUND

The importance of the Web is a *sine qua non* in modern information systems, but as can be judged by initiatives such as the *Web Accessibility Initiative* (WAI) [1], making it accessible to everyone remains a topic of considerable importance.

A recent formal investigation conducted by the *Disability Rights Commission* (DRC) [4] has brought attention to colour accessibility. The study evaluated 1000 web sites for compliance with the *World Wide Web Consortium* (W3C) *Web Content and Accessibility Guideline 1.0* (WCAG 1.0) [2] checkpoints, defined by the WAI. Notable in the results is that 81% of the web sites tested failed to satisfy the most basic WAI compliance category. *"Inappropriate use of colours and poor contrast between content and background"* accounts for 59 of the 585 usability problems, and as shown in Table 1, this problem is ranked highly across four of the six impairment groups that were included in the test.

IMPAIRMENT GROUP	INSTANCES	RANK
Partially Sighted	20	1
Dyslexia	20	3
Hearing Impaired	9	4
Physically Impaired	10	4

TABLE 1: Colour Related Accessibility Problems Experienced by Impairment Group

Surprisingly, colour related problems are dominant even in groups of people who had not reported a colour vision deficiency. This suggests either that the study group contains substantial numbers of users with undiagnosed colour vision deficiencies or that many web sites use colour in a way that is unintelligible to normal colour viewers. Regardless, strong evidence now exists indicating that colour management is a key problem for web authors [4]. In the next section we review the applicable guidelines as provided by the W3C-WAI.

3. W3C WEB CONTENT AND ACCESSIBILITY GUIDELINES

Considerable effort has been devoted to ensuring combinations of colours are safe for web users to view. The WCAG 1.0, widely regarded as the international standard for web accessibility, address colour accessibility in *Guideline 2: Don't rely on color alone*, which reads

> If color alone is used to convey information, people who cannot differentiate between certain colors and users with devices that have non-color or non-visual displays will not receive the information. When foreground and background colors are too close to the same hue, they may not provide sufficient contrast when viewed using monochrome displays or by people with different types of color deficits.

In particular, checkpoint 2.2 explicitly relates to poor contrast between foreground and background elements

> Ensure that foreground and background color combinations provide sufficient contrast when viewed by someone having color deficits or when viewed on a black and white screen.

Similarly, *Principle 1: Content must be perceivable,* in the upcoming *Web Content and Accessibility Guidelines 2.0* (WCAG 2.0) [3] recommendations tackles colour accessibility in guideline 1.4, which reads "Make it easy to distinguish foreground and information from its background."

In the following sub-sections we illustrate the associated guideline success criteria and demonstrate how they can be presented effectively so as to prevent the selection of poor colour combinations.

3.1 WCAG 1.0

The W3C's *Evaluation and Repair Tools Working Group* describes a measure of colour visibility in technique 2.2.1 of their *Techniques for Accessibility Evaluation and Repair Tools* (AERT) document [7]. The AERT document describes many techniques that web accessibility validation tools may use to evaluate and validate the conformance of HTML documents to the WCAG 1.0.

Two colours provide sufficient contrast when they differ in both colour and brightness. The colour difference D_c between colours C_1 and C_2 is

$$D_c = \sum_{i=1}^{3} \left| C_1[i] - C_2[i] \right| \tag{1}$$

where $C[i]$ is the i_{th} component (R, G, or B) of colour C. The acceptable colour difference set in the AERT requires $D_C > 500$. The brightness difference D_B between colours C_1 and C_2 is the absolute difference in luminance

$$D_b = \left| Y(C_1) - Y(C_2) \right| \tag{2}$$

where $Y(C)$ is the luminance of C in the YIQ (Luminance, In-phase, Quadrature) colour space and is calculated as

$$Y(C) = \left[0.299, 0.587, 0.114\right] \cdot \left[C_R, C_G, C_B\right]^T \tag{3}$$

The acceptable brightness difference set in the AERT requires that $D_B > 125$.

3.2 WCAG 2.0

While WCAG 1.0 was approved in May 1999 and is stable, the forthcoming WCAG 2.0 recommendations are intended to be easier to understand, use and test.

The colour visibility metric in WCAG 1.0 has been criticised as being too restrictive and has subsequently been replaced with the luminosity contrast ratio (LCR) in WCAG 2.0. The focus of the LCR is on making it easier for users to separate foreground information from the background. People with low vision often have difficulty reading text that does not contrast with its background; this can be exacerbated if the person has a colour vision deficiency that lowers the contrast even further.

Text or diagrams, and their background provide sufficient contrast when the luminosity contrast ratio, D_L between two colours C_1 and C_2

$$D_L = (L(C_1) + 0.05)/(L(C_2) + 0.05) \quad (4)$$

is at least 5:1 for Level 2 conformance or 10:1 for Level 3 conformance.
The luminosity *L(C)* is

$$L(C) = 0.2126 \times (C_R/255)^{2.2} + 0.7152 \times (C_G/255)^{2.2} + 0.0722 \times (C_B/255)^{2.2} \quad (5)$$

Luminosity values can range from 0 (black) to 1 (white), and luminosity contrast ratios are in the range [1-21]. The gamma and RGB coefficients are derived from the standard default colour space for the Internet (sRGB).

The luminosity contrast ratio is not yet a W3C recommendation but rather a suggestion by the WAI WCAG 2.0 working group to help determine whether or not the contrast between two colours can be read by people with colour blindness or other visual impairments.

Figure 1 shows these guidelines in operation[1]. WCAG 2.0 is less restrictive than WCAG 1.0 which is shown by Figure 1(h) having a smaller unacceptable zone (black coloured) than Figure 1(f) or 1(g). In all cases the unacceptable colours correspond highly to saturated colours in the swatch (shown in Figure 1(a)) which is because here we are measuring differences between the highly saturated colours in Figure 1(a) and black.

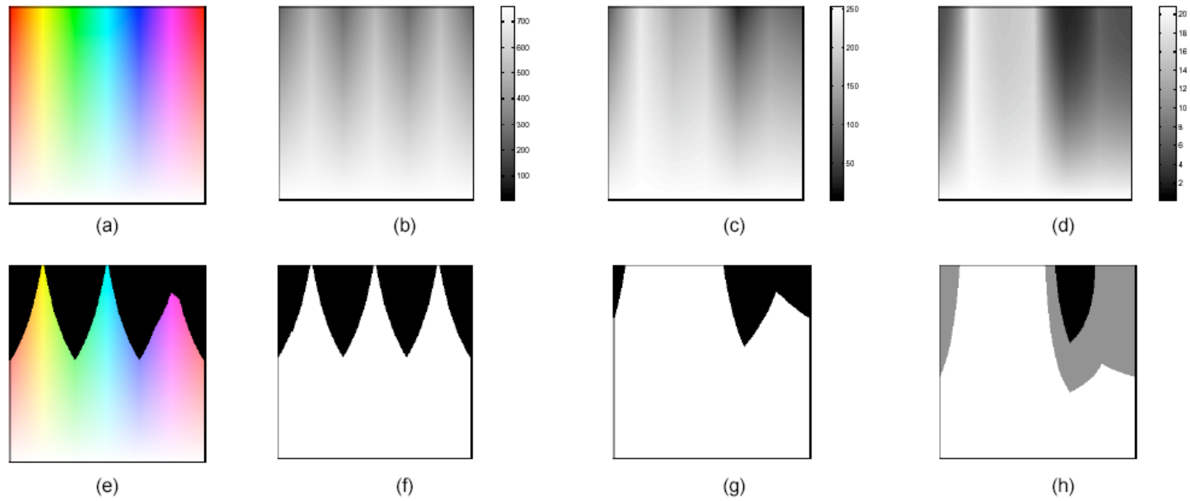


FIGURE 1: A colour swatch (a) with varying hue and saturation, and constant (maximum) brightness, some associated differences (b-d) and some visibility thresholds (e-h). (b) is D_C where C_1 = *[0, 0, 0]* and C_2 is taken from swatch (a), (c) is D_B (same colours), and (d) is D_L, (e) is the combinations of D_C and D_B via WCAG 1.0 and also the acceptable gamut, (f) shows in white $D_C > 500$, (g) shows $D_B > 125$ and (h) shows $D_L > 10$ (in white) and $D_L > 5$ (white and grey).

4. SAFE COLOUR SELECTION

Figure 2 shows our colour selection tool. A designer selects a background colour, which when fixed, indicates the range of acceptable foreground colours. The colour swatch is masked to highlight acceptable colour combinations.

We have extended the interface to allow designers to view the colours as seen by a colour vision deficient (CVD) viewer by integrating models of CVD [8]. These models are described in [9], where they were used to re-colour graphics for CVD viewers. The *Colour View* drop-down option allows the panel to be re-coloured for a particular class of CVD viewer.

[1] Caution: Figures in this paper are difficult to interpret in greyscale, see http://www.cmp.uea.ac.uk/~laj/HCI06 for colour figures.

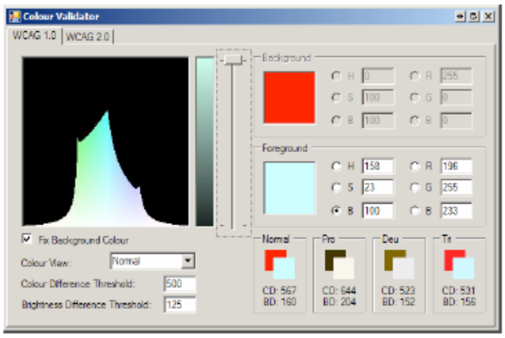


FIGURE 2: Shows part of the interface of our new colour selection tool. On the right hand side are controls for selecting the colours under consideration. On the left is a mask showing the acceptable colours for C_2 (foreground) when C_1 (background) = [255, 0, 0] via WCAG 1.0. The threshold values for the mask can be adapted to accommodate any modifications in future W3C recommendations. The two tabs allow a choice between the two sets of guidelines.

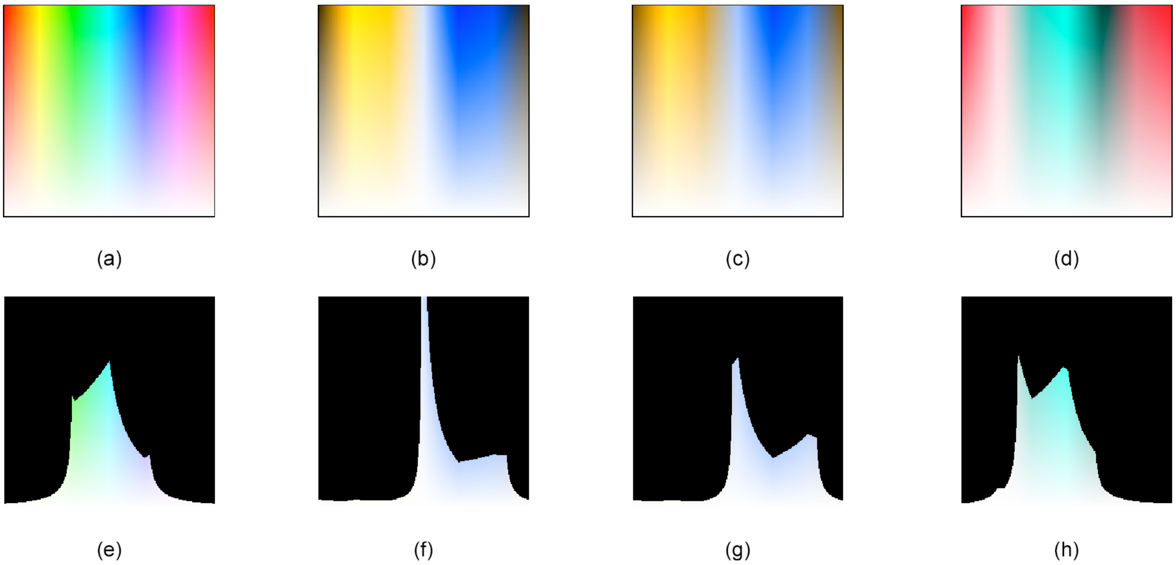


FIGURE 3: A colour swatch (a) as seen by a full colour observer and the CVD simulated versions showing how it appears to (b) protanopes, (c) deuteranopes, and (d) tritanopes. Figures (e-h) show the masks highlighting acceptable combinations D_C and D_B via WCAG 1.0 for C_2 when C_1 = [255, 0, 0] for the original and each type of CVD respectively.

Figure 3 shows the reduced gamut experienced by users with the three main types of colour blindness: protanopia, deuteranopia and tritanopia [8]. The acceptable colour regions in Figure 3(f), (g) and (h) now have considerably reduced gamut's thus emphasising the criticality of correct colour choice in these cases.

5. CONCLUSIONS AND FUTURE WORK

In this article, we proposed a novel colour selection tool to facilitate the selection of accessible colour. The tool integrates W3C evaluation criteria and hence reduces the complexity of validating colours for W3C compliance. We extended the tool to incorporate models of CVD thus allowing a web author to select colours suitable for CVD viewers.

Integrating colour accessibility into the authoring process simplifies the task of selecting accessible colour, enables non-specialists with limited knowledge to avoid poor colour combinations, and raises colour accessibility awareness. Future work needs to evaluate the practicality of using W3C success criteria for validating colours for CVD viewers.

REFERENCES

[1] W3C (2006) Web Accessibility Initiative. http://www.w3.org/WAI/
[2] W3C (1999) Web Content and Accessibility Guidelines 1.0. http://www.w3.org/TR/WAI-WEBCONTENT/
[3] W3C (2006) Web Content and Accessibility Guidelines 2.0. http://www.w3.org/TR/WCAG20/
[4] Disability Rights Commission (2004) The Web: Access and Inclusion for Disabled people. http://www.drc-gb.org/publicationsandreports/2.pdf

[5] Web Accessibility Tools Consortium (WAT-C) 2005 http://www.wat-c.org/tools/CCA/1.1/
[6] Law, Jacko, Edwards, (2005) Programmer-Focused Website Accessibility Evaluations, in Proceedings of Assets'05 (Baltimore, MD, Oct, 2005), ACM Press.
[7] W3C (2000) Techniques For Accessibility Evaluation And Repair Tools. http://www.w3.org/TR/AERT
[8] Brettel H., Viénot F., Mollon J. D., (1997) Computerized simulation of color appearance for dichromats. Journal of the Optical Society of America A, 14, 2647-2655.
[9] Jefferson L., Harvey R., (2005) Colour interfaces for colour blind computer users, in Proceedings of HCI 2005, Vol. 2, British Computer Society, 133-137.

The Role of Shame, Guilt and Embarrassment in Online Social Dilemmas

Asimina Vasalou
Intelligent Systems and Networks Group
Electrical and Electronic Engineering
Imperial College London, SW7 2BT, UK
a.vasalou@imperial.ac.uk

Adam Joinson
Institute of Educational Technology
The Open University
Milton Keynes, MK7 6AA, UK
a.n.joinson@open.ac.uk

Jeremy Pitt
Intelligent Systems and Networks Group
Electrical and Electronic Engineering
Imperial College London, SW7 2BT, UK
j.pitt@imperial.ac.uk

The self-conscious emotions of shame, guilt and embarrassment are known for regulating human societies by (1) encouraging the wrongdoer to further comply and (2) extending reparation to the one damaged. Self-awareness is a requisite for the experience of self-conscious emotions. In this paper, we hypothesise that low self-awareness online deprives an offender of the emotional consequences that usually follow a norm violation. Therefore, the aforementioned pro-social benefits of self-conscious emotions are not made possible. We test this hypothesis in a study during which online offenders were assigned to either high or low self-awareness conditions. The results show that high self-aware participants, in contrast to low self-aware participants, experience more self-conscious emotions, collaborate more when given a second opportunity and apologise more frequently.

Self-awareness, norm compliance, uninhibited behaviour, self-conscious emotions, computer-mediated communication

1. INTRODUCTION

Computer-mediated communication (CMC) has brought together people of different origins, religions and socio-economic backgrounds. Within the richness of this online communication, several theories have developed to explain online behaviour. For example, online users have been shown to perceive others in an exaggerated positive light therefore constructing their experience as 'hyperpersonal' [12]. Their online partners appear to them more attractive, more affectionate and more personable. Furthermore, context cues that signal one's belongingness to a particular social group, prime online users to follow the social conventions of that group. This has been shown to encourage behaviours that are in accord with the group's norms [8]. A central tenet of these theories is that anonymity increases attraction to others [12] and enhances identification with the group [8].

Despite the promise of CMC theories depicting the importance of online anonymity, instances of uninhibited behaviours under anonymous conditions persist. Pseudonyms are very often used as a vehicle for deception (e.g. fraud, impersonation, trolling) [4]. Indeed, next to the 'hyperpersonal' experience that may develop in CMC, there is also the possibility of 'hypernegative' behaviour. This side of CMC has been underrepresented in the most recent research to-date [12]. The present paper takes a step in addressing this gap. We depart from current CMC theories to offer and test a novel explanation for online norm violations. In the following section we argue that low self-awareness online decreases the emotional consequences that usually follow norm violations i.e. shame, guilt and embarrassment. As a result of this, we hypothesise that the pro-social benefits of those self-conscious emotions, namely behavioural compliance and reparation, break down. We then test this proposal in an online social dilemma game setting during which high and low self-aware participants were given incentives to defect. This paper ends with a discussion which generalises our results to one-off interactions in CMC.

2. A POSSIBLE EXPLANATION FOR UNINHIBITED BEHAVIOUR ONLINE: 'SELF-CONSCIOUS EMOTIONS'

Human societies offer a set of norms that their members are expected to adhere to. When a member of the community violates a norm, there are practical consequences that follow (e.g. banishment from the community), as well as emotional consequences (i.e. shame, guilt and embarrassment). Upon experiencing self-conscious emotions, people have been known to remedy their actions [11] and to pacify the offended party with apologies [1]. Even more, the experience of self-conscious emotions in the long term reinforces affiliation to those norms and encourages future compliance [11]. Therefore, self-conscious emotions play an important role both in regulating our everyday interactions and also alleviating interactions that have been disrupted.

Shame, guilt and embarrassment have been repeatedly shown to be more acute when self-awareness is increased, while in the absence of self-awareness the same emotions are experienced more weakly [11]. In a recent effort to construct a theoretical model of self-conscious emotions [11], objective self-awareness was framed at the forefront of the emotional experience. In this conceptualization, attentional focus on the self ("I") is required to instantiate self-representations (e.g. "I am a good student"). These representations are then compared against the event taking place. Incongruence between a self-representation (e.g. I am a good student) and the event taking place (e.g. getting a low grade) leads to the experience of negative self-conscious emotions. This framework indicates that without self-focused attention, one cannot attend to the event taking place and therefore instantiate a self representation in order for an appraisal of the violation to begin. Self-focused attention, also known as objective self-awareness has been described as the awareness directed inwards towards the self rather than outwards towards the environment also known as subjective self-awareness [3]. In laboratory settings, objective self-awareness has been mostly induced with mirrors which enhance self-focused attention. Objective self-awareness has also been increased by focusing the offender's attention back on his/her behaviour [2].

Contrary to what happens in face-to-face communication, in CMC, the salience of self-awareness that is otherwise created by physical and social cues (e.g. facial and corporal expressions) is reduced [7], [8]. Therefore, online, one would also expect the emotional consequences that usually follow norm violations to be experienced less acutely, offering fewer incentives for compliance. This statement is indirectly supported when observing online behaviour: one of the less desirable effects of computer-mediated communication has been the emergence of uninhibited and uncontrolled behaviour, in many cases shockingly unlike what is practised face-to-face (e.g. [5]). To address this problem, it is tempting to propose the use of stable identities where accountability would be increased by linking a norm violation to the perpetrator's real identity. However, as discussed earlier, research has shown that the inherent anonymity and physical disembodiment online, offers many benefits (e.g. [8], [10]). Consequently, in increasing behaviour regulation online, there is a need to find alternative solutions that may heighten self-awareness but at the same time will not compromise the benefits of anonymity.

3. STUDY

The study conducted was a between subjects experiment design testing self-awareness (low self-awareness vs. high self-awareness) as a factor that modulates self-conscious emotions, behaviour regulation and apology attempts following a norm violation. Participants were given incentives to 'defect' in a social dilemma game. This kind of game originated from behavioural economics [6] but has also been used in CMC [9] to better understand users' motivations during trust decisions. A participant of the game has two opposing choices (1) he can betray the trust of his opponent in order to gain more money or (2) split the gains in half, be fair and gain less money. We expected 'defectors' (those who betrayed their partners' trust) in the high self-awareness condition to experience more self-conscious emotions at the onset of the game (*H1*). As a result of this, they should be motivated to apologise or attempt to justify their actions to the offended party (*H2*). Given a second round of the game, defectors who were made high self-aware in the first round, should be motivated to rectify their behaviour by sharing more money with their opponent (*H3*). This increase in collaboration during the second round was expected to decrease both apology attempts and self-reports of the emotions experienced (*H4*).

3.1 Participants and Procedure

Fifty-six participants took part in total. However, forty participants' data was used as those participants defected and were useful to the objectives of this study. All participants were undergraduate students between the ages of 18 and 24. Proficiency in English was a requirement for attendance. Upon arriving, participants were led to a quiet room and seated by a computer. Each participant took part in the trust game [6] using a web-based application. The instructions of the game as given to participants follow.

"In this game there are two players. One of you will start off with 150 decision points and the other with 100 decision points. The player who holds 150 decision points makes the first move (first mover). The first mover can transfer 50 decision points to the other player (second mover) or keep the 50 decision points and cash in. If the first mover cashes in, this round of the game ends. If the first mover transfers 50 decision points to the second mover, the 50 points are automatically multiplied by 6 so that the decision points the second mover receives are 300. The second mover now has to decide whether to transfer part of the 300 decision points to the first mover. In this game the second mover can transfer increments of 25 decision points starting from 0 to 300. Once the second mover makes this move, the round of the game ends.

You will be randomly assigned to first or second mover but once your role is decided you are assigned to it permanently for all the rounds you play. Depending on a random draw, you may play 1 to 3 rounds of the game. You will know whether you proceed to the next round at the end of each round. You will be playing with a different player in each round."

For believability purposes, participants thought they were allocated randomly to either the first or second mover role. However, participants were always given the second mover role. In each round, participants were led to believe they were playing with a new player who was being debriefed at another UK university. The first mover was simulated in the application, always appearing to trust the participant by transferring the initial 50 points. A new non-gender indicative pseudonym was displayed in each round to represent the first mover. We 'reset' participants' interaction with the assignment of a new opponent in every new round as we wanted to control for a reputation effect that could bias the players' choices. A defection in the game constituted as transferring less than half of the amount gained i.e. less than 150 points. The game was set up so that participants could play 2 rounds. However, to provide a sense of continuity in the game, participants were led to believe they would play 3 rounds.

3.2 Experimental Design

The forty subjects of the study were randomly assigned to one of two conditions, a low self-awareness condition and a high self-awareness condition.

In the *low self-awareness condition*, participants were led to the experiment room which was faintly lit. Participants were instructed to construct a pseudonym for their participation. Low lighting and anonymity together have been used in previous CMC studies to decrease self-awareness [7].

Objective self-awareness has been increased in the lab in numerous ways, including showing the reflection of participants' image in a mirror so that the focus of self is temporal and by refocusing participants' attention on their norm violation with a reminder cue following the violation. In this paper, we focused on the latter method to construct a new self-awareness mechanism for CMC. In a study conducted by Beer et al [2] participants reported no emotion due to their inappropriate behaviour. Objective self-awareness was then increased by showing participants videos of their improper conduct therefore cuing them on their conduct. Following these retrospective viewing sessions, participants reported embarrassment. In a similar way, we propose the use of a 'text mechanism' that will act as an attention device bringing the actor's attention back to the self and the action. Hence, an actor who has violated a norm will be notified through the interface of his/her norm violation. We believe that such a mechanism will refocus the actor's attention on the anti-normative act so that an appraisal of the event can occur and therefore the possibility of experiencing a self-conscious emotion. In the *high self-awareness* condition, participants were treated identically to those in the low self-awareness condition. However, when a participant transferred less than 150 points to the first mover, the self-awareness mechanism appeared on screen. Participants read the following message: "[Participant ID], you shared less than half of your gains with [first mover ID] although your gains were made thanks to [first mover ID]'s initial donation." After participants received the self-awareness mechanism, they could not modify their money transfer decision in that round.

Three measures were collected during the study:

- *Self-conscious emotions reports*: At the end of each round, participants reported the intensity of the emotions they had experienced while playing the game. To conceal the purpose of the study, eight emotions were presented ranging from positive emotions e.g. joy to negative emotions e.g. anger. Amongst those were shame, guilt and embarrassment. Each emotion was rated from high (5) to low (1). We summed up the self-conscious emotions of shame, guilt and embarrassment into one score for each of the two rounds.
- *Apology reports*: Following each round, participants were requested to type a final message to their remote opponent in response to the question "Would you like to type a final message to [first mover ID]?" The first author and a naïve expert both rated the texts on two measures: their apologetic and their excusatory tone from high (5) to low (1). The two measures were summed up to form one score for each of the two rounds. Disagreements between raters were resolved through discussion.
- *Money transfer*: The participant's money transfer to his/her opponent was recorded in each round. The transfer was an amount between 0 and 300 and was measured in units of 25 (e.g. 100 points equalled to four units).

3.4 Results

Self-awareness manipulation check

At the end of each round of the game, participants completed a short questionnaire used in past CMC experiments as a measure of self-awareness [7]. Two questionnaire items measured objective self-awareness from high (5) to low (1): "In this experiment I've generally been very aware of myself, my own perspective and attitudes" and "Rather than thinking about myself in this experiment, my mind has been distracted by my task and what is going on around me" (reverse scored). Participants' self-awareness scores in the two rounds were summed up to one total self-awareness score. A t-test showed that participants reported significantly higher scores of self-awareness in the high self-awareness condition t (39) = -2.610, $p<0.01$. Therefore, the self-awareness manipulation worked.

Analysis

A Kolmogorov-Smirnov test showed that the data was not normally distributed. We therefore chose a Mann-Whitney non-parametric test for the analysis. Emotion, money transfer and apology for round 1 and round 2 were the dependent measures. Table 1 presents the descriptive statistics.

	Low self-awareness condition (N=20)		**High self-awareness condition (N=20)**	
	Mean	Std. Deviation	Mean	Std. Deviation
Money transfer round 1	2.25	1.71	2.45	1.986
Money transfer round 2	1.75	1.74	3.25	2.593
Emotion report round 1	4.50	1.76	6.70	2.155
Emotion report round 2	4.40	1.72	5.30	2.386
Apology round 1	2.78	1.61	4.60	2.644
Apology round 2	2.88	2.07	3.88	2.025

TABLE 1: Descriptive Statistics

In round 1, participants of the high self-awareness condition reported higher self-conscious emotions (U=85.00, $p<0.01$). Our first hypothesis (H1) was confirmed: higher self-awareness increased participants' emotions after defecting. No difference was found in the money transfer metric in round 1 (U=185.00, $p>0.05$). Participants made equal transfers which indicate that the defectors were equal in both conditions. In support for H2, high self-aware participants made more efforts to apologise and justify their actions (U=116.00, $p<0.05$). In round 2, the pattern of results was inverted. In agreement with H3, participants who had been made self-aware in the first round, transferred more money in round 2 (U=132.500, $p<0.05$). As a result of this, those participants reported less emotions (U=152.00, $p>0.05$). However, contrary to our prediction, high self-aware participants still apologised more often, although the effect was weaker (U=152.00, $p<0.05$). Conclusively, in the second round, high self-aware participants who had experienced more emotions in the first round transferred more money. Participants' adherence to the norm explains their lower reports of self-conscious emotions and the weaker apology effect.

SUMMARY

While the integration of computer-mediated communication in everyday life is inevitable and irrefutably beneficial, understanding the causes of online norm violations and proposing new ways to strengthen behaviour regulation are both relevant and important topics in CMC and to the prevalent themes of this conference.

The game scenario chosen for this study can be generalised to one-off online interactions where participants do not have expectations of further interacting with their online partner in the future. In fact, according to [12], low interaction expectancy lowers attraction to others which may also explain reduced inhibition online. In this paper, we demonstrated that during one-off online social dilemmas, being fair and cooperating with strangers may be hindered by low self-awareness. We began with the premise that low self-awareness online reduces an offender's experience of shame, guilt or embarrassment. We claimed that this effect in return may reduce the pro-social derivatives of those emotions: behaviour regulation (i.e. compliance in future interactions) and affiliation attempts (i.e. apologies). We described a text mechanism for increasing self-awareness which we tested in an online social dilemma game setting. Our hypotheses were all confirmed: self-aware participants reported experiencing significantly more self-conscious emotions in the first round of the game that low self-aware participants. Similarly, those participants apologised more frequently. In the second round, self-aware participants reported fewer emotions, apologised in lesser degrees and transferred significantly more money to their new online opponent. Therefore, the experience of self-conscious emotions in the first round of the game reinforced participants' affiliation to the norm in the second round. Given our results, we can claim that in those one-off encounters during which anti-normative behaviour can be increased, the elicitation of self-conscious emotions (e.g. by means of self-awareness) may serve as a check that ultimately promotes behaviour regulation and repair within CMC.

Acknowledgements
Special thanks to Tobias Regner, Jens Riegelsberger and Tanja Bänziger for many interesting discussions. Etienne Roesch is gratefully acknowledged for his insights on this work. We thank Lidia Oshlyansky for her assistance. This research was funded by the HUMAINE IST Framework VI Network of Excellence.

REFERENCES

[1] Apsler, R. (1975). Effects of embarrassment on behavior toward others. Journal of Personality and Social Psychology, **32**, 145-153.
[2] Beer, J. S., Roberts, N. A., Werner, K. H., Scabini, D., Levenson, R. W. and Knight, R. T. (2002) The role of self-conscious emotion in self-regulation. Proceedings of the Annual Convention of the Society for Personality and Social Psychology, Savannah, Georgia.
[3] Carver, CS (1979), "A cybernetic model of self-attention processes," Journal of Personality and Social Psychology, **37, 8**, 1251-1281.
[4] Donath, J. (1998) Identity and Deception in the Virtual Community. In M. Smith and P. Kollock (eds.), Communities in Cyberspace. Routledge, London.
[5] Grady, D. (1998) Faking pain and suffering on the Internet. The New York Times.
[6] Hopfensitz, A. (2006) The Role of Affect in Reciprocity and Risk Taking. PhD thesis, Tinbergen institute, Amsterdam.
[7] Joinson, A. N. (2001) Self-disclosure in computer-mediated communication: The role of self-awareness and visual anonymity. European Journal of Social Psychology, **31**, 177-192.
[8] Postmes, T., Spears, R., Sakhel, K. and De Groot, D. (2001) Social influence in computer-mediated communication: The effects of anonymity on group behavior. Personality and Social Psychology Bulletin, **27**, 1243-1254.
[9] Riegelsberger, J., Sasse, M. A. and J. McCarthy (2003). The Researcher's Dilemma: Evaluating Trust in Computer Mediated Communications. International Journal of Human Computer Studies, **58**, 6, 759-781
[10] Stritzke, W. G. K., Nguyen, A., and Durkin, K. (2004) Shyness and Computer-Mediated Communication: A Self-Presentational Theory Perspective. Media Psychology, **6**, **1**, 1-22
[11] Tracy, J. L., and Robins, R. W. (2004). Putting the self into self-conscious emotions: A theoretical model. Psychological Inquiry, **15**, 103-125.
[12] Walther, J. B. and Parks, M. R. (2002) Cues filtered out, cues filtered in: Computer-mediated communication and relationships. In Knapp M. L. and Daly J. A. (eds.), Handbook of interpersonal communication. Sage, Thousand Oaks, CA.

Managing Online Music: Attitudes, Playlists, Mood and Colour

Russell Beale and Michael Voong
Advanced Interaction Group, School of Computer Science, University of Birmingham
Edgbaston, Birmingham, B15 2TT UK

The movement of music from physical discs to digital resources managed on a computer has had an effect on the listening habits of users. We present the results of a survey to quantify these changes, and identify problems with the creation and management of playlists, a mechanism which users use to organize and guide their listening experiences. We present the results of an experiment to determine the effectiveness of associating colour with music. We demonstrate that there is a reasonable degree of consistency between user's associations of colour and music, despite their personal views that everyone will be different.

MP3, online music, interfaces, search & retrieval, colour tagging

1. INTRODUCTION

The migration of music from analogue to digital format, from vinyl LPs to resin CDs, and then from CD to MP3, from the disk and onto the computer, has caused a change in the way we think about our music and the soundtrack to our lives. The move onto the computer has connected music and listeners to the wider world in the same way that the web connected individual users to a wider community for information sharing. We can now access music from around the world, downloading it from one machine to another with minimal effort. This sharing of music came to the fore with Napster and other peer-to-peer file sharing services, rapidly closed down by the combined forces of legal argument and commercial pressure from the copyright holders and record companies. This has been reborn in the form of online music stores, the most high profile of which is Apple's iTunes store. A lot of research and commercialization has been invested in digital rights management to provide some control of the dissemination of music, but much less effort has gone into understanding how users now integrate their music into their lives and interact with it, now that it is in this new format and stored in a different location.

For many users, the change of media for their music has altered their use of music as well. Informal discussions with people suggest that new ways of interacting with their music have developed. Some comment that they have rediscovered much of their music collection since the random access provided by the players makes later tracks as accessible as the early ones. When a CD is played, it is often played from the beginning in order, and poor or disappointing tracks can lead to the CD being replaced by another before it finishes, ensuring that the later tracks remain unplayed. Computer media players democratize the process, since they provide simple direct access to any track. A typical music management system and player is shown in Figure 1. Tracks are tagged with supplementary information, which provides information on the song title, duration, artist, album, genre, user rating, album artwork, and (not shown here) a count of how often it has been played, the date and time it was last played, the date added to the collection, track number on the album. It is also possible to display technical information about the file size, sampling bit rate and sample rate as well. Most computer music systems allow music tracks to be tagged in this way, and most utilize the Gracenote database [1], which provides an online service to provide key information about tracks, artist and title information automatically. This tagging allows music to be ordered by more than just artist or album title – all songs by an artist, irrespective of the album on which they appear, can be collected together, or all songs ordered alphabetically, for example. This process moves the atomic unit of music from the album back to the individual track.

Users are therefore able to play tracks in sequence, where the sequence is an ordering on any of the tags associated with tracks: typically artist's name, album title or song title are used. However, the interface supports rapid movement through the library of material, and any individual track can be selected with scrolling and a click of the mouse. Supporting this random access mode of playing is the 'shuffle' option, in which tracks from the user's whole collection are played in a random order. This provides them with an experience of their music much like a radio station, in that they have no direct control over what is played next though their general taste is catered for, much as they can choose the radio station they listen to.

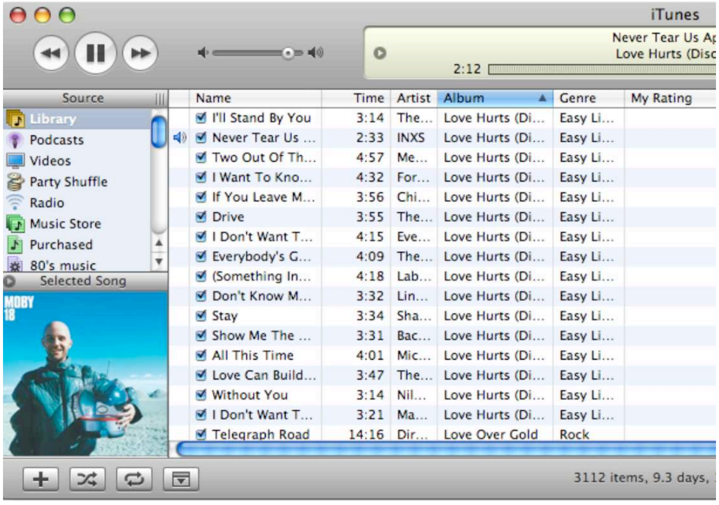


FIGURE1. iTunes music player interface

2. USER RESEARCH AND ANALYSIS

In order to understand more about users and their new relationship with their music, we conducted an online survey to understand more about users and their attitudes. The survey was publicised to students, family and friends, and to readers of the usabilitynews.com website who are HCI practitioners, researchers and those with a general interest. 56 users took part over a period of two months. The raw data is presented in Table 1.

Age		**Different Genres**		**Playback Method**	
Under 16	1	Under 10	22	Create playlists	18
Between 17 and 20	22	Between 10 and 20	30	Play Individual Songs	12
Between 21 and 30	27	Over 20	10	Play whole album	19
Between 31 and 40	7			Shuffle	8
Over 40	5	**New Tracks**		"Smart Shuffle"	1
		Daily	5		
Media Player		Weekly	19	**Collection Size**	
Foobar 2000	2	Monthly	23	Less Than 100 Songs	8
iTunes	23	Rarely	15	Several Hundred or More	19
MusicMatch Jukebox	1			A Thousand or More	28
Winamp 3/5	9	**Ratio Tagged/Untagged**		Tens of Thousands	7
Winamp Classic (2)	1	1/4 or Less	5		
Windows Media Player	18	More Than 1/4, less than 1/2	10	**iTunes Ratings Usefulness**	
SysTrayPlay	1	More Than 1/2	20	Useful	16
The Core Media Player	1	All or Virtually All	27	Useless	36
Real Player	1			Very Useful	4
Sonic Stage	1				

TABLE 1: Survey results into computer-based music management

Most of the users were between 17 and 30 (94%), reflecting the generally youthful takeup of this form of music usage. Users tend to have significant music collections: 56% of users have the equivalent of more than 100 CDs worth. This survey was self-selecting, though not targeted at music listeners or mp3 experts in particular, but since it discussed music and music players, it is likely that this will bias the survey towards users with an interest in the topic and hence likely to have larger collections. A variety of media players were used by users, with iTunes (39%) leading Windows Media Player (30%) and Winamp (16%). Most users tag their songs: 76% of users have at least half of their collection tagged.

Our survey suggests that conventional approaches to listening to music are still prevalent: 33% play the whole album as their prime approach: 21% play individual songs. However, 14% experience their songs in a random order by default, whilst almost 1/3 of users create playlists. A playlist is a user-defined collection and ordering of music, often chosen because it reflects a mood or style of music. The iPod shuffle device, a highly portable solid state mp3/aac player, only allows users to play songs in playlist order, or random order; it deliberately eschews conventional listening habits. Playlists are created by selecting tracks and adding them to a list: any track can be added to any playlist - tracks can be by one artist, or on a particular topic, but appear to often be attuned to a certain mood or emotion. For example, a commonly reported experience is creating a 'party' playlist or one for driving to, or a mellow chill out one for late evenings: one user requested an

> "auto-playlist where you don't have to make the playlist, it selects the songs in terms of moods, style, genre, not artists or album. Having music picked for you in whatever mood you are in."

Another commented that

> "the music player should learn my music taste and help me select the right music for the moment. Something fresh in the morning, something smooth in the evening ..."

with another adding

> "I would like the automated playlist generator to work better. It would be interesting to see if there is any logic to my listening habits i.e. I tend to only listen to rock in the morning and rap & hip hop in the evening."

In an attempt to automate the process of creating playlists, iTunes has a facility for users to rate songs, from zero stars to five stars. However, nearly 2/3 of users do not find this a particularly useful feature. This is probably because it offers a rating of the quality of a song, which is not the basis on which users chose to listen to their music. Digital music makes it easy to skip tracks that you really do not like, leaving tracks that you are prepared to listen to. Without any further categorization, this is not sufficient to allow a useful clustering of songs. It is unlikely that a user will want to listen to a whole set of only "average" songs, for example. This is supported by a user who stated (sic):

> "the 5 star rating system is kinda flawed becuase it doesn't really take mood or genre into account, it'll tell me if i think a track is good or not but that doesn't necessarily mean i'll feel like listening to a given 5* track right now. I find that i have masses of mp3s but half of them arn't really organised well in genre etc, either becuase they were mislabled or not labled at all and so i find it difficult to find a whole bunch of tracks easily that i can just start playing immediately in itunes, instead i have to resort to making my own playlists which is time consuming or playing a particular album which constrains me to listening to one particular artist at a time The user interface of itunes is generally good though, except that if i begin making a lot of playlists then i end up with a list thats only organised alphabetically whereas the individual tracks can be organised with artist, genre... But bear in mind i only use playlists excessively becuase itunes doesn't have a good enough inbuilt way of organising my music by mood"

This user research has identified that the management of online music is something that requires better user support: the evidence of the survey shows that users most often use playlists for their music and that they have substantial collections of music. It shows that the current user-selectable tags (the rating system) is regarded as inadequate, and users would like a way of managing their music by mood.

3. COLOUR AND MUSIC

A number of approaches to automatically generating playlists exist. Platt et al. [2] provide a system that learns from a selection of existing songs, though this requires the user to select a set of songs initially, and this doesn't solve the problem of providing a list to suit a mood: instead it focuses on providing alternatives to identified songs. Alternative approaches [3, 4] use analysis of existing track characteristics (timbre, dynamics etc.) to generate similar songs from those currently selected. Our system design differs, and is focused on the fact that many playlists seemed to be created to suit a certain mood, and so we should provide a tagging mechanism that was related to mood [5]. It was decided to use colour to provide the expressivity and simplicity required. Colour is often associated with emotion and moods, supported by existing psychological theories [6-8]), and it provides a natural and simple way of expressing something about a song [9]. It is true that different cultures can have radically different interpretations of colours, but a common western set is shown in Table 2.

We decided to undertake a trial to determine general reactions to colour from music lovers. Respondents to the original survey were contacted to ask them to take part, and were asked to listen to a specific track and then to indicate the colour that they most associated with that track. Users were also asked to comment on the association between colours and music tracks.

Users felt that associating colours with music was interesting but fraught with problems: most felt that users would be hugely inconsistent in their colour associations, and that this would cause problems in sharing information between users. Typical comments include (sic):

1. "Its sometimes hard deciding which colour appeals to me when listening to some tracks either because a colour does not come up in my mind or there is more than one colour."
2. "Music generates such a complex serious of thoughts, images and feelings that to associate a song with a colour may be oversimplifying the process."
3. "it's an interesting idea, but can't see how it will work if everyone has such different opinions on music/colours"
4. "using the idea of colour isnt good for there would be too many opinions for everyone has different ideas or thoughts about how a song should match colour"

Red: energy, danger, courage, power, passion and desire	Light red: joy, sensitivity and love Pink: romance, friendship, feminine qualities Dark red: willpower, anger leadership, wrath Brown: stability, masculine qualities
Orange: sunshine, enthusiasm, fascination, happiness, creativity, stimulation	Dark orange: deceit and distrust Red-orange: pleasure, aggressions, action Gold: wisdom, wealth
Yellow: joy and cheerfulness	Dull yellow: decay, sickness and jealousy Light yellow: freshness and joy
Green: suggest harmony and symbolises growth. Strongly associated with safety and stability.	Yellow-green: seasickness, cowardice, discord and jealousy Dark green: ambition, greed and jealousy Aqua: emotional healing and protection Olive green: traditional colour of peace
Blue: depth, trust and truth	Light blue: health, healing, tranquillity and softness Dark blue: knowledge and power
Purple: ambition, wisdom, independence, creativity, mystery and magic.	Lavender: femininity, nostalgia, romantic feelings Dark purple: gloom, sadness, frustration
White: cleanliness, safety, goodness, innocence, virginity	
Black: power, elegance, death, evil, mystery	
	(taken from [8])

TABLE 2: Colours and their typical associations

User 1 and 2 both think that there is no repeatable association between colour and music: users 3 and 4 have fewer problems on a personal level but wonder about the consistency across different users. The second issue is less of a problem for our design, in that we are interested in providing a tool for individual users, and if they have some form of colour – music mapping that works for them, then we can use this to classify music appropriately. User 5 offers a deeper insight into their understanding of the problems:

5. "In the past, I've organized my CD collection by the color of the CD spine (e.g. all the whites together). I believe we often locate objects and information by color more than we know or might like to admit. Lucy Suchman found that law office workers file case folders by color, for instance. As a librarian, I saw people going re-locating books by their color. For personal organization of music, this could be a useful idea. However, I strongly believe that there is no intrinsic property in music that suggests color. This would be subjective and change even for a given person from situation to situation. Any association of color is external to the music itself and imposed upon it from without."

The trial we undertook presented five mp3 files to 57 users and asked them to associate the song with one of ten given colours (black, brown, grey, orange, purple, red, green, yellow, blue, white).

value	song 1	song 2	song 3	song 4	song 5
black	2	1	15	1	2
brown	6	4	6	0	2
grey	12	4	10	1	15
orange	3	12	1	12	1
purple	5	22	1	4	4
red	4	4	19	17	2
green	9	4	1	3	6
yellow	4	2	4	11	4
blue	11	4	0	7	11
white	1	0	0	1	10

TABLE 3: Frequency of choice of colour for the 5 songs

The results in Table 3 show variation in the strength of choice of the most representative colour, but for all of them we can reject the null hypothesis, that there is no correlation between colour and the music, at the p=0.01 level. The results indicate a reasonable degree of agreement between different users about the colour to associate with a given track, though the effect is stronger for common genres of music than uncommon ones. For our design concept, all that matters is that the user is self-consistent in their choice, though these results suggest that the effects are strong enough to support a shared recommendation system which recommends songs based on their colours as a measure of similarity. These results suggest that one or two colours are strongly associated with songs, and that there is a diverse colour palette in play.

REFERENCES

1. Gracenote, *http://www.gracenote.com/.* 2006.
2. Platt, J., et al. *Learning a Gaussian Process Prior for Automatically Generating Music Playlists*. in *Neural Information Processing Systems (NIPS)*. 2001.
3. Logan, B. *Content-based playlist generation: Exploratory experiments*. in *Proceedings of ISMIR*. 2002.
4. Vignoli, F. and S. Pauws. *A Music Retrieval System Based on User-Driven Similarity and its Evaluation*. in *Proceedings of International Conference on Music Information Retrieval (ISMIR)*. 2005.
5. Feng, Y., Y. Zhuang, and Y. Pan. *Music Information Retrieval by Detecting Mood via Computational Media Aesthetics*. in *WI '03: Proceedings of the IEEE/WIC International Conference on Web Intelligence*. 2003: IEEE.
6. Levkowitz, H., *Color theory and modeling for computer graphics, visualization, and multimedia applications*. 1997, Boston, USA: Kluwer Academic Publishers. 219.
7. Campbell, I.G., *Basal Emotional Patterns Expressible in Music.* The American Journal of Psychology, 1942. 55(1): p. 1-17.
8. *Color Wheel Pro - Color meaning*. 2005.http://www.color-wheel-pro.com/color-meaning.html
9. Kawanobe, M., M. Kameda, and M. Miyahara. *Corresponding affect between music and color*. in *Systems, Man and Cybernetics, 2003. IEEE International Conference on*. 2003.

Mining Users' Preferences in an Interactive Multimedia Learning System: a Human Factor Perspective

Kyriacos A. Chrysostomou, Sherry Y. Chen and Xiaohui Liu
School of Information Systems, Computing and Mathematics, Brunel University, Uxbridge, UB8 3PH, UK
{Kyriacos.Chrysostomou, Sherry.Chen, Xiaohui.Liu}@brunel.ac.uk

With the increasing popularity of the Internet, interactive multimedia learning systems are widely used in both educational and industrial settings. However, as human factors vary across users, they may prefer the design of interactive multimedia learning systems differently. To have a deep understanding of the influence of human factors, we apply a data mining approach to the assessment of users' preferences in using interactive multimedia learning systems. More specifically, the K-modes algorithm is used to group users' preferences. The results indicate that users' preferences could be divided into four groups where computer experience significantly influences their preferences. Implications for the development of interactive multimedia learning systems are also discussed.

e-Learning, interface, human factors, data mining, vlustering, K-modes

1. INTRODUCTION

One of key features of e-Learning applications is to use the capabilities of interactive multimedia, which applies various media, such as text, images, audio, animation and video [18]. With the growing popularity of the Internet, interactive multimedia learning systems have been widely adopted by both educational and industrial settings. However, much remains to be learned about how different users perceive such systems. Users are unique and have a variety of human factors that significantly influence their learning patterns [17]. In this way, each user will appreciate such media differently which will in turn determine whether they can successfully accept and use interactive multimedia learning systems [1]. Therefore to enhance the users' experience and satisfaction, human factors should be considered as an essential issue for the development of these systems.

To this end, there is a need to investigate user's preferences in using interactive multimedia learning systems in order to match system features to their diverse needs [11]. To have a deep understanding of user preferences, this paper proposes to use a data mining approach, which has the ability to discover hidden concepts and potentially useful and interesting relationships from data. There are a number of data mining approaches. Based on the nature of their information extraction, many of these approaches can be divided into three major categories: clustering, classification, and association rules [3]. Among these categories, clustering is selected because of its ability to partition data into groups (also known as clusters), which share similar characteristics [6]. In this way, we can obtain a general understanding of how different users interact with interactive multimedia learning systems.

Previous research suggests that human factors influence user preferences. For example, Mitchell, Chen and Macredie (2005) found that experts tend to have more positive perceptions to non-linear navigation than novices [13], while Roy and Chi (2003) indicated that females and males are more likely to adopt different methods for navigation [14]. Therefore, the dominant human factors of each cluster will also be examined in this study. By doing so, these human factors can be considered in the development of interactive multimedia learning systems. Accordingly, the paper is structured as follows. Section 2 describes the experiment designed to collect users' preferences and the techniques applied to analyse the data. Subsequently, the grouped users' preferences are illustrated in Section 3, where the effects of human factors are also discussed. Finally, conclusions are drawn and possibilities for future work are identified in Section 4.

2. METHODOLOGY

2.1 Empirical Study

Eighty participants took part in the experiment, during which they used two interactive multimedia learning systems only once. More specifically, all users were given System A first, followed by System B. Both learning systems shared exactly the same content and adopted a quiz-style format to deliver general knowledge questions (e.g. sport, entertainment, and history). Although these two systems shared the same content, they used different interaction styles to present their interfaces in order to identify users' preferences. System A adopted the WYSIWYG (What You See Is What You Get) interaction style, while System B used the WIMP (Windows, Icons, Menus, Pointers) environment as its interaction style. These two interaction styles were chosen because they are commonly used in the creation of e-Learning applications. In addition to interaction styles, other differences between these two interactive multimedia learning systems lie within the interface layout, button types, colour schemes, multimedia elements and menu formats. These differences (Table 1) provided the basic rationale for the design of the questionnaire. The questionnaire was used to, firstly, capture the users' demographics so as to obtain the individuals' human factors, including age, gender and computer experience. In terms of computer experience, users were asked to indicate how often they used computers and software packages. In addition, users were required to select the interface features they most favoured from the choices offered by the questionnaire, so as to identify their preferences of both interactive multimedia learning systems.

Interactive Features	System A	System B
Interface Layout	Single window	Multiple windows
Button Types	Static (which do not give an indication, i.e. a colour change, when pressed), without embedded icons	Dynamic (which change colour or form when pressed), with embedded icons
Colour Scheme	Multiple colours with the addition of effects, i.e. blending one colour into another.	Few standard colours.
Multimedia Elements	Images, graphics, audio and video	Images, graphics and audio
Menu Formats	Without drop-down menus	Drop-down menus to access the help, images, and audio

TABLE 1: The Differences Between The Two Interactive Multimedia Learning Systems

2.2 Data Analysis

2.2.1 Preprocessing of Data

The data preprocessing stage predominantly involved attribute selection. The attributes that did not relate to the users' preferences were excluded so that any deterioration with regards to the clustering of instances could be reduced [4]. For example, attributes which specifically related to the quiz, such as the type of questions or results feedback preferred by users, were excluded. As a result, the selected attributes corresponded to the different multimedia features provided by the two interactive multimedia learning systems (Table 1). Thus, the final set of attributes comprised of: (1) the layout of the interface, (2) the button type preferred by users, (3) the use of icons embedded within buttons, (4) the use of menus, and finally (5) their preferred colour scheme.

2.2.2. K-modes Algorithm

Among a plethora of clustering algorithms, the K-means algorithm is a widely known and used technique for grouping objects with similar characteristics, mainly due to its computational efficiency [9]. The K-modes algorithm is an extension of K-means algorithm, used to cluster data containing mixed numeric and categorical values [8]. The K-modes algorithm uses a simple matching dissimilarity measure to deal with categorical objects, replaces the means of clusters with modes, and uses a frequency-based method to update modes in the clustering process to minimise the clustering cost function. With these extensions, the K-modes algorithm enables the clustering of categorical data in a fashion similar to K-means. Such extensions are useful for analysing data of this study, because the data obtained through the questionnaire are categorical.

Like K-means, the K-modes algorithm requires the number of clusters (k) and the seed (s), which generates the values for the assignment of the initial cluster centres, to be fixed a priori. Since the algorithm is sensitive to how clusters are initially assigned [10], it is necessary to try different values and evaluate the results in order to find the combination that better fits the data. This is because different runs of the algorithm, i.e. changing k and s, yield different results. Consequently, different combinations of the previously mentioned attributes were used to evaluate results for the best performance of the algorithm. Having exhausted several combinations, the results showed that the algorithm produces more meaningful outcomes when k = 4 and s = 10.

3. RESULTS AND DISCUSSION

3.1. Interactive Multimedia Features

The clustering of users has shown a definitive divide between their preferences of interactive multimedia features. Table 2 illustrates the meaning of each cluster with regards to users' preferences of features found in both interactive multimedia learning systems. The chosen attributes indicate that participants are grouped according to the following trends: (1) Cluster 1: users prefer the single window interface that utilises static buttons with no embedded icons and no use of drop-down menus though they favour the colours with effects background; (2) Cluster 2: users prefer the multiple window layout, as opposed to users in Cluster 1, use dynamic buttons with embedded icons and favour drop-down menus along with the standard colour scheme; (3) Cluster 3: users synonymously favour the single window interface as in Cluster 1, prefer static buttons with embedded icons and the multicoloured background scheme, though they do not like the drop-down menus; and (4) Cluster 4: users synonymously prefer the multiple window interface with dynamic buttons that do not contain icons, do not use drop-down menus and favoured the colours with effects style.

As depicted in Figure 1, many users appear in Cluster 2 and few emerge in Cluster 4. The main differences between Cluster 2 and Cluster 4 lie within the use of icons, drop-down menus, and colour schemes. The users in the former prefer dynamic buttons with embedded icons, drop-down menus, and like the standard colour scheme, while those in the latter favour dynamic buttons without icons, dislike drop-down menus, and prefer the colours with effects format.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Interface Layout	Single Window	Multiple Windows	Single Window	Multiple Windows
RECOMMENDATION: ***utton Type***	Static	Dynamic	Static	Dynamic
Use of Icons	No	Yes	Yes	No
Use of Menus	No	Yes	No	No
Colour Scheme	Colours with Effects Scheme	Standard Colour Scheme format	Multiple Colour Scheme	Colours with Effects Scheme

TABLE 2: The Differences Between The Four Clusters

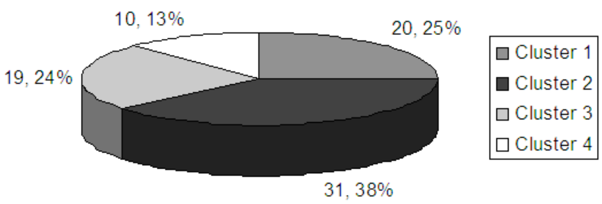


FIGURE 1: The Number of Users in Each Cluster

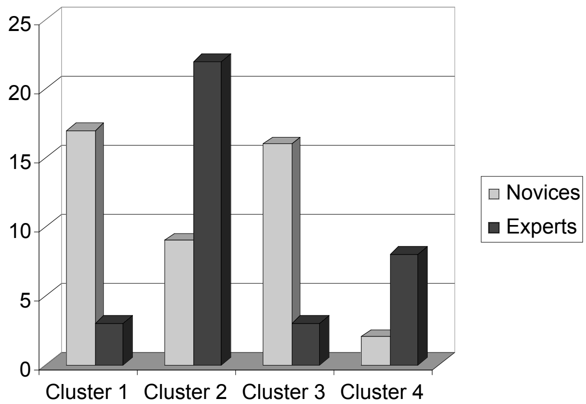


FIGURE 2: Levels of Computer Experience in Each Cluster

As aforementioned, Cluster 2 consists of more users (N=31, 38%) than the remaining three clusters. The key difference between Cluster 2 and the other clusters is that the users in this cluster favour a single colour scheme. This suggests that the single colour scheme is most popular with users. In contrast, multiple colours and colours with effects are found to be less popular with users. These results are compatible with the cognitive load theory [19], which suggests the focus of an instructional module must be the instruction itself. Information that is adjunct to the instruction must be designed to minimise cognitive load [5]. In this study, the single colour scheme may in turn increase the users' concentration on the instruction itself. On the other hand, multiple colours and colours with effects could promote distraction and unnecessary clutter to the users' mind and exacerbate cognitive load as well as associated mental energy. This may explain why most users prefer the single colour scheme, instead of multiple ones and colours with effects.

Moreover, Cluster 4 has the least number of users (N=10, 13%). Users in Cluster 4 prefer the multiple window layout, dynamic buttons, and the colours with effects scheme. This may suggest that the integration of these interactive multimedia features offered users a pleasant visual display with multiple colour presentation.

By examining the demographic information of these ten users, it is interesting to see that all are females. A noticeable difference between females and males is that female users particularly favoured appealing images as a means of presenting information [12]. This difference may be able to explain the reason why female users prefer interactive multimedia learning systems with attractive visual displays, as illustrated by the features they chose in Cluster 4.

3.2. The Effects of Human Factors

It is surprising to find that computer experience is the dominant human factor of each cluster. Figure 2 illustrates the proportion of experts and novices within each cluster. The majority of experts (N=30, 83%) appeared in Cluster 2 and Cluster 4 whereas novices (N=33, 75%) mainly emerged in Cluster 1 and Cluster 3. A detailed discussion is presented below.

3.2.1. Interface Layout

A difference between Clusters 2/4 and Clusters 1/3 is that the users in Clusters 2/4 like multiple windows (referring to System B) while those in Clusters 1/3 prefer a single window layout (referring to System A). It suggests that novices prefer a single window layout whereas experts prefer the multiple window layouts. In other words, one's computer experience can dramatically affect his/her preferences of interface layout. Such findings echo previous work by Smith (1999), who found that computer experience may in turn affect one's motivation when using the system [16]. Novices have greater difficulty in assimilating interfaces they have previously never seen due to their lack of prior experience. Thus, they prefer an interface that does not require them to heavily rely on their prior expertise or knowledge. Single window design is one type of such interface. On the contrary, experts have gained more theoretical insight and a number of guiding principles to infer ambiguous computer scenarios [1], so they feel comfortable to interact with more complicated interface layouts, e.g. multiple windows.

3.2.2. Navigation Tools

Navigation buttons and main menus are two major navigation tools utilised in interactive multimedia learning systems. According to our results, the preference of navigation buttons is another difference between Clusters 2/4 and Clusters 1/3. The static button is favoured by the users in Clusters 1/3 whereas the dynamic button is preferred by the users in Cluster 2/4. The other difference between Clusters 2/4 and Clusters 1/3 is the use of drop-down menus. The drop-down menus are favoured by the users in Clusters 2/4, instead of those in Clusters 1/3. In other words, the majority of experts favour using dynamic buttons and the drop-down menus; while novices like static buttons and dislike drop-down menus. A possible reason for such findings may be that dynamic buttons and drop-down menus belong to more advanced interactive multimedia features, which are beneficial to experts. However, these features may not be useful to novices who have little or no formal training and experience [15]. These results are in line with those of Hasan, which found that individuals perceive themselves at a disadvantage when they do not have sufficient computer experience to enable them to complete their task [7]. It may be due to the fact that novices exhibit higher levels of anxiety [2] and such anxiety can affect the way in which they used the interactive multimedia learning system.

4. CONCLUSIONS

The study presented in this paper adopted a data mining approach to uncover relationships between human factors and users' preferences in interactive multimedia learning systems. Overall results revealed a prominent divide between diverse types of users, as shown by their varied preferences across clusters where computer experience had considerable effects on preferences. As discussed in Section 3, these results reinforce the findings of previous research that indicated experts and novices favoured different types of interface features, especially in interactive multimedia learning systems. Our findings as well as those in previous research can be integrated together for developing personalised interactive multimedia learning systems that can accommodate the needs and preferences of different users. Nonetheless, this is only a small-scale study. Further work needs to be undertaken with a larger sample to provide additional evidence. Moreover, it would be interesting to analyse users' preferences with other clustering techniques, such as hierarchical clustering or fuzzy clustering. Finally, the findings of such studies could be integrated to build robust user models for the development of effective personalised interactive multimedia learning systems.

REFERENCES

[1] Antonietti, A. and Giorgetti, M., (2006) Teachers' beliefs about learning from multimedia, *Comp. in Human Behavior*, 22, (2), 267-282.
[2] Beckers, J., et al., (2006) The influence of computer anxiety on experienced computer users while performing complex computer tasks, *Comp. in Human Behavior,* 22, (3), 456-466.
[3] Chen, S. Y. and Liu, X., (2004) The Contribution of Data Mining to Information Science. *J. of Information Science*. 30, (6), 550-558.
[4] Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P., (1996) From Data Mining to Knowledge Discovery in Databases: A Survey, *AI Magazine*, 17, (3), 37-54.
[5] Feinberg, S. and Murphy, M., (2000) Applying cognitive load theory to the design of web-based instruction. *In Proc. of IEEE Professional Communication Society international Professional Communication Conference and Proc. of the 18th Annual ACM International Conference on Computer Documentation: Technology & Teamwork (Cambridge, Massachusetts, pg. 353-360).*
[6] Hand D.J., Mannila H., and Smyth P., (2001) *Principles of data mining*, MIT Press, 293-296.
[7] Hasan, B., (2003) The influence of specific computer experiences on computer self-efficacy beliefs, *Comp. in Human Behavior*, 19, (4), 443–450.
[8] Huang, Z, (1998) Extensions to the K-means algorithm for clustering large data sets with categorical values", Data Mining and Knowledge Discovery, 2, 283-304.
[9] Jain, A. K., Murty, M. N., and Flynn, P. J., (1999) Data clustering: A review, *ACM Computer Surveys*, 31, (3), 264-323.
[10] Khan, S. and Ahmad, A., (2004) Cluster center initialization algorithm for *K*-means clustering, *Pattern Recognition Letters,* 25, (11), 1293-1302.
[11] Lohr, L. L., (2000) Designing the instructional interface, *Comp. in Human Behavior*, 16, (2), 161–182.
[12] Miller, H. and Arnold, J., (2000) Gender and Web home pages, *Comp. & Education*, 34, (3/4), 335-339.
[13] Mitchell, T. J. F., Chen, S. Y. and Macredie, R. D., (2005) Hypermedia Learning and Prior Knowledge: Domain Expertise vs. System Expertise, *J. of Computer Assisted Learning*, 21, (1), 53-64.
[14] Roy, M. and Chi, M. T. C., (2003) Gender differences in patterns of searching the web, *J. of Educational Computing Research*, 29, (3), 335-348.
[15] Simmons, P. E., and Lunetta, V. N., (1993) Problem-solving behaviors during a genetics computer simulation: beyond the expert/novice dichotomy, *J. of Research in Science Teaching*, 30, (2), 153-173.
[16] Smith, B., et al., (1999) A review of the construct of Computer Experience. *Comp. in Human Behavior* 15, (2), 227-242.
[17] Southwell, B. G., et al, (2004) Translating user control availability into perception: the moderating role of prior experience, *Computers in Human Behavior,* 2004, Available online at: [Accessed: 12th December 2005]
[18] Sun, P.-C. and Cheng, H. K., (In Press) The design of instructional multimedia in e-Learning: A Media Richness Theory-based approach, *Computers & Education*.
[19] Sweller, J., van Merriënboer, J. J. G. and Paas, F., (1998) Cognitive architecture and instructional design, *Educational Psychology Review,* 10, (3), 251–296.

Student Attitude to Adaptive Testing

Mariana Lilley, Trevor Barker
University of Hertfordshire, School of Computer Science, College Lane, Herts AL10 9AB, United Kingdom
M.Lilley@herts.ac.uk, T.1.Barker@herts.ac.uk

A computer-adaptive test (CAT) is a computer-assisted assessment application in which the test dynamically adapts itself to the proficiency level of individual students. To enhance student engagement, CAT software applications aim to provide students with tasks that are sufficiently challenging, and yet not so difficult that could lead to boredom or frustration. The CAT prototype introduced here comprised a graphical user interface, a database of questions and an adaptive algorithm based on the Three-Parameter Logistic Model from Item Response Theory. A group of 76 Computer Science undergraduate students participated in a summative and a formative assessment session using our CAT prototype. At the end of each session, participants were asked to rate the level of difficulty of the overall test from 1 (very easy) to 5 (very difficult). The perceived level of difficulty of the test and the CAT scores obtained by the participants were subjected to Spearman's rank order correlations. Findings from this statistical analysis suggest that participants' perceptions of difficulty were not related either to performance or to the type of test undertaken.

Computerised adaptive testing, adaptive interfaces, intelligent tutoring systems, student engagement

1. INTRODUCTION

In Higher Education today, increasing reliance is being placed upon the use of online systems. Often these are used to manage learning, present information and test students in an entirely undifferentiated way, all users having exactly the same view of the system. With the development of increasingly large and complex computer applications and greater diversity in student groups, consideration of individual differences has become an important issue in designing usable and useful applications [1][13][14].

User modelling is a technique that is often employed to this end, allowing users to perform tasks and interact with systems differentially, depending on some feature of their personalities, abilities, preferences or performance. Intelligent software applications that adapt to their users based on user models have been gaining rapidly in importance within the human-computer interaction field. Examples of such adaptive interfaces include systems that help users to filter web query results [15], recommendation systems that help users to make choices [7], intelligent tutoring system applications that adapt to the knowledge of individual users [4]. Computer-adaptive tests (CATs) are a further example of a software application within the intelligent interfaces domain [6], and this type of adaptive software is the focus of this paper.

A CAT is a computer-assisted assessment application in which the test dynamically adapts itself to the proficiency level of individual students. The underlying idea is that students may become bored and demotivated with questions that are too easy for them. Similarly, students may become frustrated and disengaged with questions that are above their proficiency level and therefore too difficult. To enhance student engagement, CAT software applications aim to provide students with tasks (i.e. questions) that are sufficiently challenging, and yet not so difficult that could lead to boredom or frustration.

Brusilovsky [5] cites CATs as one of the elements of a paradigm shift within educational software development, from "one size fits all" to one capable of offering higher levels of interaction and personalisation. Despite a growing body of research in CAT, student attitude toward the CAT approach has been under represented in the literature [6].

2. SOFTWARE PROTOTYPE OVERVIEW

The CAT software prototype introduced here comprised a graphical user interface, a database of questions and an adaptive algorithm. The output from the CAT prototype is a student profile, based on cognitive skills and proficiency levels per topic area within the subject domain. This profile is used to generate a student model which, in turn, is employed to provide feedback to students [10].

2.1 Adaptive Algorithm

CATs are based on Item Response Theory (IRT). IRT is a family of mathematical functions that attempts to predict the probability of a user successfully completing a task or, more specifically, answering a question correctly. In the software prototype introduced here, the adaptive algorithm was based on the Three-Parameter Logistic (3-PL) model from IRT. A fuller account of IRT can be found in Lord [12] and Wainer [16].
A typical CAT starts with a question of average difficulty. In general terms, a correct response will cause a more difficult question to be administered next. Conversely, an incorrect response will cause a less difficult question to follow.

2.2 Student model

One of the central elements of the 3-PL Model is the level of difficulty of the task being performed by the user. Indeed important assumptions of the model include the need to provide a database of questions that is accurately ranked according to question difficulty.

Two approaches were then used to achieve the task's difficulty estimate. Initially, subject experts were then used to rank the questions in order of difficulty, based upon their experience of the subject domain and Bloom's taxonomy of cognitive skills [3]. This was undertaken independently by two experts who later met to agree a final ranking. An initial value of the difficulty parameter was established in this way ranging from -3 to +3, as shown in Table 1.

At the end of each assessment session, user performance was employed to adapt and update the values of the difficulty parameter for each question, thus providing a more objective measure of difficulty. Questions that were answered correctly more often had their difficulty level lowered and vice versa. A study of the difficulty levels in the question database over time was able to show that the initial expert calibration was accurate and fair, and that the post-test adaptation was effective in establishing accurate difficulty parameters [2].

Difficulty range	Cognitive skill	Skill involved
-3 <= b <=-1	Remember	Ability to recall taught material
-1 < b < 1	Understand	Ability to interpret and/or translate taught material
1 <= b <= 3	Apply	Ability to apply taught material to novel situations

TABLE 1: Cognitive skills

As illustrated in Figure 1, the questions in the database were also classified according to topic area (e.g. ADO.NET). This was deemed necessary for the generation of a student profile differentiated according to topic area within the subject domain.

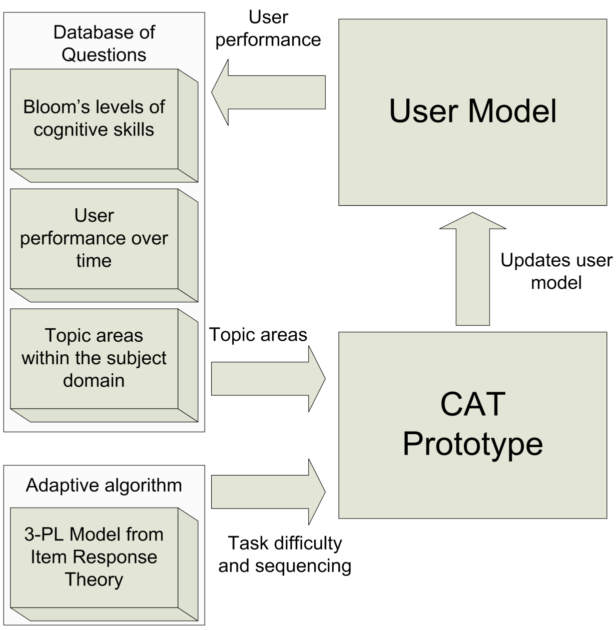


FIGURE 1: The diagram illustrates how the CAT software prototype works

3. THE STUDY

As part of their regular assessment for a programming module, a group of 76 Computer Science undergraduates participated in two assessment sessions using our CAT software prototype. The first assessment session was formative and therefore the scores obtained by the participants did not count towards their final grade. The second assessment session was summative. In both cases, participants had 40 minutes to answer 40 objective questions within the Visual Basic.NET subject domain.

Table 2 shows a summary of their assessment performance. In Table 2, the value for the proficiency level ranged from -3 (lowest) to +3 (highest).

Assessment	Proficiency level Mean	Std. Deviation
Formative	-0.03	1.02
Summative	0.21	1.42

TABLE 2: Summary of participant performance (N=76)

At the end of each test, students were asked to rate the difficulty of the test that they have just taken from 1 (very easy) to 5 (very difficult). The test difficulty mean, as perceived by the participants, was 3.53 (SD=0.64, N=76) for the formative test and 3.46 (SD=0.59, N=76) for the summative one. Their ratings are illustrated in Table 3.

Assessment	1 (Very easy)	2 (Easy)	3 (Just right)	4 (Difficult)	5 (Very difficult)
Formative	0	2	36	34	4
Summative	0	2	39	33	2

TABLE 3: Level of difficulty of the test as perceived by the participants (N=76)

It was important to investigate whether or not the correlation between participants' performance and their perceptions on the level of difficulty of the overall test was statistically significant. Such statistical analysis is the focus of the next section of this paper.

4. STATISTICAL ANALYSES

Participants' results and their perception of the test's difficulty were subjected to Spearman's rank order correlations and Kruskal-Wallis tests. In addition, a paired-samples t-test was used to examine any significance of differences in their means between formative and summative assessment sessions.

4.1 Formative assessment session

No statistically significant correlation was found between the participants' proficiency levels and the test's difficulty rating (rs = -0.165, Sig. 2-tailed = 0.155, N = 76). The data gathered in this study was also subjected to a Kruskal-Wallis test (Chi-Square = 3.591, df = 2, Asymp. Sig. = 0.166). As can be seen from Table 4, the participants' performance on the formative test had no effect on the perceived difficulty of the test.

Group	N	Mean Rank
Low performing participants	25	44.54
Average performing participants	26	34.58
High performing participants	25	36.54

TABLE 4: Perceived level of difficulty of the formative assessment test, according to performance (N=76)

4.2 Summative assessment session

The findings for the summative assessment session were in line with those in the formative one. No statistically significant correlation was found between the participants' proficiency levels and the test's difficulty rating (r_s = -0.025, Sig. 2-tailed = 0.829, N = 76). The Kruskal-Wallis test mean ranks are shown in Table 5 (Chi-Square = 4.336, df = 2, Asymp. Sig. = 0.114). These results were taken to indicate that the participants' performance on the summative test had no effect on the perceived test difficulty.

Group	N	Mean Rank
Low performing participants	26	41.27
Average performing participants	24	31.65
High performing participants	26	42.06

TABLE 5: Perceived level of difficulty of the summative assessment test, according to performance (N=76)

4.3 Comparisons between formative and summative assessment sessions

The absence of statistically significant relationship between performance on the test and perceived test difficulty is an interesting finding. The perception of the difficulty of a test might be expected to relate in some way to performance. Although it is difficult to be certain of a reason for our finding, it is consistent with our view that the test we provided was effective in establishing the appropriate level of difficulty for individual learners. This is of particular importance, since one of the goals of our CAT prototype was to provide individual users with tasks that are engaging, rather than tasks that are uninteresting or frustrating. We argue that establishing an appropriate level is necessary, though of course not sufficient, to achieve this objective.

A paired-samples t-test was used to examine any significant differences in the means for the perceived level of difficulty obtained for the two assessment sessions (i.e. formative and summative). No statistically significant difference was found (t = 0.799, df = 75, Sig. 2-tailed = 0.427). A paired-samples t-test was also used to examine any significant differences in the means for the proficiency level obtained for both assessment sessions. This test showed statistically significant differences between proficiency level means (t = 0-2.112, df= 75, Sig. 2-tailed = 0.038).

Whilst there were no statistically significant differences in the perceived level of difficulty means, there were statistically significant differences in the proficiency level means. The proficiency level mean for the summative test (Mean = 0.21, SD = 1.42, N = 76) was higher than the formative one (Mean = -0.03, SD = 1.02, N = 76). The fact that students are more likely to revise for a summative test than for a formative one could explain the difference in performance. It is also possible that students adopt different strategies and they are more meticulous when taking summative tests. Another possibility is that the formative test had a positive effect on students' preparation for the summative test.

5. SUMMARY AND DISCUSSION

Students differ in their interests, career intentions, ways of learning, cognitive styles and proficiency levels within the subject domain. The CAT software prototype introduced aims to enhance engagement by tailoring the level of difficulty of the tasks to the proficiency level of individual students. In previous work [8][9], we have shown that the CAT approach is effective when measuring student proficiency levels. However, this work has focused on the use of the CAT approach in a summative assessment context. It was therefore important to investigate student attitude in a formative assessment scenario. This work offers a new perspective by providing a point of comparison between student attitude to a CAT implementation of formative and summative assessments.

The results reported here suggest that:

- There was no difference in the perception of difficulty of the CAT approach that could be ascribed to the type of test, i.e. formative or summative
- There was no difference in the perception of difficulty of the CAT approach that could be ascribed to test performance

These findings support our view that the goal of providing students with a set of engaging tasks, regardless of their proficiency levels within the subject domain, was supported by the establishment of a tailored test for each student based on the inherent adaptive characteristics of the CAT approach.

In a CAT, the level of difficulty of the test relates to student proficiency levels rather than directly to the intrinsic difficulty of the learning objectives being tested. It may be argued that this approach does not provide an objective estimate of student proficiency level. In our earlier research we were able to show that the CAT approach provided a fair, accurate and reproducible measure of proficiency level [8][9][11]. It is true to say however, that the interpretation of the results of CAT assessments requires further research in the future.

A further future direction of our work will be to investigate how the student model introduced here can be used in a wider educational context, such as the delivery of learning materials. One potential application would be a system that attempted to predict learning resources that a student may be interested in or may need to revise, based on the information in the student profile.

REFERENCES

[1] Barker T., Jones S., Britton C., Messer D. (2002) The use of a cooperative student model of learner characteristics to configure a multimedia application. User Modelling and User Adapted Interaction 12 (2/3), pp. 207-241.

[2] Barker T., Lilley M., Britton C. (2006) A student model based on computer adaptive testing to provide automated feedback: The calibration of questions. Paper accepted and in press, 13th International Conference of the Association for Learning Technology.
[3] Bloom B. S. (1956) Taxonomy of educational objectives. Handbook 1, Cognitive domain: the classification of educational goals, London, Longman.
[4] Brusilovsky P. (2001) Adaptive Hypermedia. User Modeling and User-Adapted Interaction 11, pp. 87-110.
[5] Brusilovsky P. (2004) Knowledge Tree: A Distributed Architecture for Adaptive E-Learning. Proceedings of 13th World Wide Web Conference, May 17-22, New York, pp. 104-113.
[6] Jettmar E., Nass C. (2002) Adaptive testing: effects on user performance. Proceedings of SIGCHI Conference on Human Factors in Computing Systems: Changing our world, changing ourselves. Minneapolis, pp. 129-134.
[7] Langley P., Thompson C., Elio R., Haddadi A. (1999). An adaptive conversational interface for destination advice. Proceedings of Third International Workshop on Cooperative Information Agents, Uppsala, Lecture Notes in Artificial Intelligence (LNAI) 1652, pp. 347-364.
[8] Lilley M., Barker T. (2003) Comparison between computer-adaptive testing and other assessment methods: An empirical study. Research Proceedings of 10th Association for Learning and Teaching Conference, Sheffield, pp. 249-258.
[9] Lilley M., Barker T. (2004) A Computer-Adaptive Test That Facilitates the Modification of Previously Entered Responses: An Empirical Study. Proceedings of 7th Intelligent Tutoring Systems International Conference, Maceió, 30 August - 3 September, Lecture Notes in Computer Science 3220, pp. 22-33.
[10] Lilley M., Barker T. (2005) The use of Item Response Theory in the development and application of a user model for automatic feedback: A case study. Proceedings of 19th British HCI Group Annual Conference, Edinburgh, 5-9 September, pp. 99-104.
[11] Lilley M., Barker T., Britton C. (2004) The development and evaluation of a software prototype for computer adaptive testing. Computers & Education Journal 43(1-2), pp. 109-123.
[12] Lord F. M. (1980) Applications of Item Response Theory to practical testing problems. Lawrence Erlbaum Associates, New Jersey.
[13] Sutcliffe A., De Angeli A. (2005) Assessing interaction styles in web user interfaces. Proceedings of Human Computer Interaction INTERACT 2005, Rome, 12-16 September, Lecture Notes in Computer Science 3585, pp. 405-415.
[14] Uruchrutu E., MacKinnon L., Rist R. (2005) Designing the learning interface using cognitive styles. Proceedings of 19th British HCI Group Annual Conference, Edinburgh, 5-9 September, pp. 105-110.
[15] Vouros G. (2001). Conceptual modelling of multimedia objects for user-tailored information presentations. Journal of Applied Artificial Intelligence 15, pp. 521-559.
[16] Wainer H. (2000) Computerized Adaptive Testing (A Primer). Lawrence Erlbaum Associates, New Jersey.

A Cautionary Tale: Hofstede's VSM Revisited

Lidia Oshlyansky
University of Wales
Swansea
Swansea, SA2 8PP
lidiaosh@gmail.com

Paul Cairns
UCL Interaction Centre
London WC1E 7DP
p.cairns@ucl.ac.uk

Harold Thimbleby
University of Wales
Swansea
Swansea, SA2 8PP
h.w.thimbleby@swansea.ac.uk

Hofstede's cultural model has been widely used to make sense of the differences seen in cross-cultural HCI research. Hofstede's Value Survey Module (VSM) and the cultural indices it produces are well known in the HCI community. This paper reports on a recent re-examination of the VSM (specifically, VSM94) in nine countries. Contrary to expectations, the data collected did not replicate Hofstede's findings. Factor analysis reveals that VSM questions are not resulting in robust, replicable factors. We discuss possible issues in the method of data collection, but given that our method follows that of many other similar studies, our results suggest that the VSM should be interpreted with caution, particularly if it is to be used to adapt user interfaces for different cultures.

VSM,cCulture, cross-cultural HCI, cultural indices

1. HOFSTEDE'S VSM IN HCI

The field of HCI has investigated the effects of culture on user interface design, use, adoption and usability of various technologies. Often culture is studied within the framework of Hofstede's classic cultural model [1, 2], which is comprised of five dimensions captured by his Value Survey Module (VSM). The VSM dimensions have been widely used as a framework to provide design guidelines, for instance for cross-cultural website development [3]. Other researchers have attempted to test the guidelines based on Hofstede's model [3] by matching subjects' cultural profile to the cultural profile of the website, though with mixed results [4]. Hofstede's model has also been used as a framework to explain cultural differences in existing website designs [5]. The Hofstede model is widely cited; indeed, out of 6 papers in the 2005 British HCI conference discussing culture, 4 cited Hofstede.

Despite its wide use and acceptance, Hofstede's work has been criticised for its lack of insight into the richness and depth of culture. It has been suggested that a more qualitative or activity theory approach would be more appropriate [6, 7]. Hofstede's research has also been criticised because it focuses on national cultures. The methodology used in his research has also been brought into question as, for example, his entire sample was drawn from IBM employees [8].

Nonetheless, given the extensive use of the Hofstede model, it may still have value in capturing some aspects of usability cross-culturally. We therefore performed a large-scale study to put the implications of the VSM on a quantitative footing by studying it in relation to other commonly used instruments in HCI. Our approach fits with Hofstede's view of using the VSM to conduct secondary research [9]. This paper focuses solely on the data collected from the VSM; a larger study and analysis is still underway. Surprisingly, our data does not replicate any of the rankings of countries against the different dimensions found in Hofstede's previous work. Principal component analysis was used to analyse our VSM data in more detail and none of Hofstede's original dimensions are reflected robustly in our data. As the method of our study follows that of Hofstede's and others in using the VSM to measure culture, our results suggest that the cultural dimensions are not as widely generalizable as previous literature would suggest.

2. METHOD OF DATA COLLECTION

The aim of the study was to collect VSM data (amongst others) from students in eleven countries, which would replicate Hofstede's work and also update it with a younger, different demographic than the IBM employees originally used. Hofstede's VSM94 was translated into six languages: Arabic (Saudi Arabian) Czech, Dutch, French, Greek and Malay. Each translation was done by at least two bilingual speakers using the back-translation process to ensure as much accuracy as possible [10]. The questionnaires were distributed to university students in the Czech Republic, France, Greece, India, Malaysia, Netherlands, New Zealand, Saudi Arabia, South Africa, the United Kingdom, and to the United States. In all countries students were recruited from diverse faculties including: Humanities, Science, Health Science, Medicine, Engineering and Computer Science, Business and Economics. No sample was represented by less than 5 academic disciplines.

A total of 1428 questionnaires were returned. To ensure a reasonably representative sample for each country, only those countries with close to 100 (or more) usable questionnaires were included in the analysis. Only those questionnaires that were completed by natives of the country being sampled were used in the analysis – “native” being defined as having been wholly educated in and living in the country being sampled. Unfortunately, despite our best efforts, insufficient responses were received from France and the Netherlands, leaving nine countries in our final sample: Czech Republic, Greece, India, Malaysia, New Zealand, Saudi Arabia, South Africa, the United Kingdom and the United States. The highest number of questionnaires was returned from Malaysia (168) and the lowest from Saudi Arabia (91). The final sample consisted of 519 men and 489 women (72 respondents did not give their gender). The average age was 23.4 years.

3. VSM SCORES AND RANKINGS

For each of the nine countries, the scores for each dimension were calculated using the formulas provided by Hofstede [9]. These scores are shown in Table 1. In brackets below each score is the ranking for the given dimension obtained in our study, followed by the ranking based on Hofstede’s scores. Saudi Arabia has no rankings as it was not studied as an individual country by Hofstede; in addition there is only limited information available on the Time Orientation dimension.

	Czech Republic	**Greece**	**India**	**Malaysia**	**New Zealand**	**United States**	**South Africa**	**United Kingdom**	**Saudi Arabia**	**Kendall rank correlation**
Power Distance	35.0 (2/4)	50.4 (1/3)	31.4 (3/2)	23.4 (5/1)	20.9 (7/8)	20.8 (8/6)	23.2 (6/5)	30.2 (4/7)	29.1	0.38
Uncertainty Avoidance	83.2 (7/2)	113.4 (1/1)	97.0 (3/6)	97.4 (2/7)	81.3 (8/3)	83.9 (5/5)	89.3 (4/3)	83.7 (6/8)	93.6	-0.36
Individualism	85.1 (6/5)	94.0 (4/7)	78.0 (8/6)	80.4 (7/8)	96.4 (3/3)	97.3 (2/1)	87.3 (5/4)	103.3 (1/2)	88.2	0.64
Masculinity	17.0 (6/5)	45.2 (2/5)	49..0 (1/7)	33.0 (4/8)	11.2 (7/4)	31.8 (5/3)	34.7 (3/2)	7.3 (8/1)	49.9	-.40
Time Orientation	54.0 (4/5)	56.2 (1/)	42.2 (8/1)	54.3 (3/)	51.6 (5/2)	46.2 (7/3)	48.3 (6/)	54.4 (2/4)	42.4	-0.60

TABLE 1: VSM Scores for each country with ranking and Hofstede’s original ranks and Kendall correlations

Hofstede recognises that, across studies, scores may not be exactly the same. However the relative rankings of countries should remain reasonably consistent. To quantify the degree of agreement, Kendall’s rank order correlation coefficient was used as it provides a value between –1 and +1 indicating the degree to which the two rankings agree on the orderings given to the different countries. Only one of these results suggests a strong agreement in ranking, namely, the ranks based on Individuality, and this is significant (p=0.026). Interestingly, the only other two correlations approaching significance, namely those for Masculinity and Time Orientation, are actually showing the *reverse* ordering from Hofstede!

The fact that one dimension does seem to carry over from Hofstede’s study to this study is in itself good, but it is unexpected that it is the only one to do so. In order to better understand the structure of our data, we performed a factor analysis on the questionnaire data; this is discussed next.

4. FACTOR ANALYSIS OF THE VSM DATA

Hofstede’s dimensions are essentially factors derived from his VSM questionnaire data. If these factors are robust, a similar analysis of our data should result in factors that closely resemble the original VSM dimensions. Hofstede [9] points out that, ideally, repeat analysis should use at least 10 countries - whereas we have only nine. If the factor structure is robust, however, then a smaller set of countries is likely to reveal similar factors, but of course these factors may be conflated to produce amalgams of the original VSM dimensions.

	1	2	3	4	5
IDV - q1	**0.42**	-0.02	-0.11	-0.06	**-0.53**
IDV - q2	**0.60**	0.08	-0.04	-0.11	-0.19
IDV - q4	**0.67**	-0.17	0.12	0.00	-0.07
IDV - q8	**0.51**	**0.33**	**-0.32**	-0.12	0.12
TO – q10	**0.53**	-0.24	**0.44**	-0.13	-0.05
TO – q12	**0.50**	-0.11	**0.49**	-0.03	-0.22
MAS – q15	-0.02	**0.62**	0.13	0.12	-0.18
MAS – q20	-0.01	0.15	**0.37**	-0.13	**0.66**
MAS – q5	**0.60**	0.15	-0.15	0.14	0.15
MAS – q7	**0.57**	0.07	-0.22	-0.17	0.16
PDI - q14	0.03	**0.50**	0.28	**-0.33**	-0.13
PDI - q17	0.24	0.02	**0.35**	**0.43**	0.21
PDI - q3	**0.55**	0.02	-0.28	0.24	0.16
PDI - q6	**0.65**	0.17	-0.14	0.05	0.21
UAI - q13	-0.12	**0.62**	0.23	-0.29	-0.08
UAI - q16	-0.20	**0.47**	-0.07	**0.34**	0.02
UAI - q18	0.02	0.18	0.11	**0.68**	-0.22
UAI - q19	0.11	-0.10	**0.54**	0.13	0.05

TABLE 2: VSM Principal Components Analysis for all countries

Principal Component Analysis was run on the questionnaire data from all nine countries with direct oblimin rotation in order to reveal the underlying simple factor structure [11]. As the samples were large (well over the suggested minimum 100 respondents suggested) the cut-off for significant loading of 0.3 was used [11]. All those variables that loaded above 0.3 or below –0.3 are shown in bold in Table 2. The left hand column lists the abbreviations for each index (IDV, Individualism; TO, Time Orientation; MAS, Masculinity; PDI, Power Distance; UAI, Uncertainty Avoidance) and the number of the question as it appears on the VSM94 (e.g., q1). The numbers in bold indicate the major constituent questions of a given factor. The expected picture from this process would be that there is an initial omnibus factor followed by separate factors that reflect the VSM dimensions or that are conflations of two or more of these dimensions (given that we have fewer than 10 countries). Whilst there does seem some sort of omnibus first factor, there is very little in the rest of the table to suggest any marked similarity between the VSM dimensions and the factor structure of our data. Oddly, Uncertainty Avoidance does not load on the first factor, and two questions from each of Masculinity and Power are also missing from the first factor. Additionally, none of the other factors strongly match with *any* of the VSM dimensions. The only possible exception to this is Time Orientation (in factor 3) but this does not match with Hofstede, as his loading coefficients reflect negative correlation between questions 10 and 12 and the Time Orientation factor, whereas factor 3 indicates positive correlation between these questions and that factor. There is a similar problem with the Individualism dimension, having a mix of positive and negative loadings in Hofstede's [9] equations but only having positive loadings in factor 1 for our data.

It is possible that the lack of expected loadings is due to some error in the way the VSM was administered. The VSM questionnaire was translated into several other languages: Arabic, Czech, Greek and Malay, and it is possible that the translated versions were not working "as they should" so two more factor analyses were conducted using data from just those countries sampled with the English version of the questionnaire. The first analysis included all the countries sampled in English: India, New Zealand, South Africa, UK and the USA. This produced results that were no better than those seen in Table 2.

The second analysis excluded India and South Africa. In both these countries English is used in teaching and business but is not always used in the home and it can often be a "second" language for many. It was considered that perhaps the questionnaire had not worked as expected because of some complication in interpreting the language used in the questionnaire in India and South Africa. In the case of South Africa it is important to state that Hofstede's sample included only white South Africans whereas our sample includes a broader cross section of the population attending university. This could also have contributed to the difference in scores seen here. Likewise, South Africa is still often said to be a country made up of different cultures, it may therefore be difficult to view it as one cohesive "culture." To rule out these issues the analysis was run again without India and South Africa. The results, as can be seen in Table 3, still do not show any of the VSM dimensions emerging as strong features in any of the factors. This suggests that whatever the problem with our use of VSM, it is not solely due to the translation process or the other issues discussed above.

	1	2	3	4	5	6
IDV - q1	**0.41**	0.03	-0.28	-0.14	**-0.49**	0.11
IDV - q2	**0.49**	0.23	0.07	-0.18	-0.12	0.19
IDV - q4	**0.65**	-0.18	0.03	-0.10	0.04	0.13
IDV - q8	**0.43**	**0.50**	-0.20	-0.17	**0.31**	0.18
TO - q10	**0.43**	**-0.49**	**0.36**	-0.01	-0.01	0.05
TO - q12	**0.45**	-0.22	**0.34**	0.03	-0.17	**0.45**
MAS - q15	0.05	**0.54**	0.21	0.17	-0.23	0.09
MAS - q20	-0.12	0.00	**0.52**	-0.22	**0.45**	0.09
MAS - q5	**0.63**	0.06	0.07	0.19	-0.08	**-0.47**
MAS - q7	**0.59**	0.13	-0.12	**-0.31**	**0.37**	0.10
PDI - q14	0.03	**0.31**	**0.46**	-0.14	**-0.44**	0.03
PDI - q17	-0.01	0.10	**0.38**	**0.37**	**0.39**	0.18
PDI - q3	**0.55**	0.04	-0.11	**0.44**	0.04	-0.27
PDI - q6	**0.63**	0.07	-0.12	0.21	0.16	-0.12
UAI - q13	0.05	**0.44**	**0.44**	**-0.34**	-0.13	**-0.32**
UAI - q16	-0.16	**0.42**	0.18	0.16	0.17	-0.21
UAI - q18	-0.03	0.12	0.16	**0.67**	-0.14	0.30
UAI - q19	0.15	**-0.49**	**0.41**	-0.04	-0.04	**-0.34**

TABLE 3: VSM Principal Components Analysis for primary English speaking countries

5. OTHER POSSIBLE CAUSES FOR UNEXPECTED VSM LOADINGS

It is difficult to say why the VSM dimensions do not emerge in our dataset. It seems safe to rule out translation as a problem since the English-language only samples do not show any better factor loadings than the mixed language set. It may be possible that this is due to some other aspect of the data set, such as age. Age does influence some VSM dimensions, UAI and MAS for example [2]. Education level could also be contributing to some of the data peculiarity. Hofstede himself [9] cites work done to correct for education level for the various dimension scores, but he does not mention this as a problem for factoring the raw data. Hence neither of these issues can be completely ruled out. Also, Hofstede [9] suggests that the VSM94, which was used for this research, had not been employed enough to prove its validity without a doubt. Likewise, he suggests (in [9]) that at least 10 countries be used for a truly reliable cross-country test, whereas the present research only has 9. However, as discussed earlier, some semblance of the VSM dimensions would have been expected to emerge in our data set. Possibly the most significant issue trying to replicate the factors by which Hofstede originated his dimensions is that we are not using the original questionnaires Hofstede used, but the version he now suggests, the VSM94.

6. CONCLUSIONS

This study conducted a straightforward administration of the VSM94 questionnaire, yet it was unable to replicate any of Hofstede's original dimensional distinctions, with the possible exception of individuality. Factor analysis suggested that the VSM dimensions had very little explanatory power in explaining the structure of our large dataset. Our results call into question what validity the VSM model has both in itself and as a tool for understanding the design of user interfaces for different cultures, although it remains a useful shared language for discussion. The educational background and age group of the participants are possible issues in our study, but we are confident that the translation of the questionnaire is not an issue. And even if it were, the question still is *when* can VSM be used as a reliable indicator of cultural differences? At best, we can suggest that the VSM was measuring something, but what it measures is as yet unclear. Future researchers will need to demonstrate that the VSM data is appropriate for their purposes, that it can be used to explain the cultural differences being considered for design.

Obviously more research is needed in this area to determine what the cultural factors are that are relevant to good interaction design, and which would support HCI research in general. It is not enough to observe the difference in interactions and interfaces from one culture to the next and to explain these observations in the light of a cultural model, especially one that was not originally intended for this purpose. Further research will need to explore the VSM and other cultural models with other populations and in other contexts to understand exactly what help cultural models can be to HCI.

ACKNOWLEDGEMENTS

Lidia Oshlyansky did the survey and analysis and is the main author of this paper. She has an Overseas Research Student Award and a PhD scholarship from Swansea University. Harold Thimbleby has a Royal Society-Research Merit Award. These sources of funding are gratefully acknowledged in the support of this research. Also, a great amount of gratitude goes to all the people who helped translate the questionnaires and gather the data in the countries sampled, unfortunately too numerous to list here; thank you!

REFERENCES

[1] Hofstede, G. (1991). *Culture and Organizations: Software of the Mind*. McGraw-Hill,New York.
[2] Hofstede, G & Hofstede, G. J. (2005). *Culture and Organizations: Software of the Mind*. McGraw-Hill, New York.
[3] Marcus, A. & Gould E. (2000). Crosscurrents: Cultural Dimensions and Global Web User-Interface Design. *Interactions*, **7**, 33–46.
[4] Ford, G. & Gelderblom, H. (2003). Effects of Culture on Performance Achieved through the use of Human Computer Interaction. *Proceedings of the 2003 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologies on Enablement through Technology* (SAICSIT) Stellenbosch, South Africa, 4-6 October, 218–230.
[5] Gould, E. W., Zakaria, N., & Yusof, S. A. M. (2000) Applying culture to website design: a comparison of Malaysian and US websites. *Proceedings of the 18th Annual ACM International Conference on Computer documentation: technology & teamwork*. Cambridge, Massachusetts, 24–27 September, 161–171.
[6] Ratner, C. & Hui, L. (2003). Theoretical and Methodological Problems in Cross-cultural Psychology. *Journal of the Theory of Social Behavior*, **33** 67–94.
[7] Baskerville, R. (2003). Hofstede never studied culture. *Accounting, Organizations and Society*. **28**, 1–14.
[8] McSweeney, B. (2002). Hofstede's Model of National Cultural Differences and their Consequences: A Triumph of Faith - A Failure of Analysis. *Human Relations*. **55**, 89–118.
[9] Hofstede, G. (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions, and organizations across nations* (2nd edition). Thousand Oaks, California: Sage Publications, Inc.
[10] Neuman, W. L. (2000) *Social Research methods: qualitative and quantitave approaches* 4th edition. Allyn and Bacon, Boston.
[11] Kline, P. (2002). *An Easy Guide to Factor Analysis*. Routledge, New York.

Designing Educational Software to Enhance the Creative Learning experience: an Integrative Framework

Sylvia M. Truman and Paul Mulholland
Knowledge Media Institute, the Open University, Milton Keynes, UK MK7 6AA
s.m.truman@open.ac.uk; p.mulholland@open.ac.uk

When learning about a subject area in the classroom, the acquisition of a set of core concepts is often perceived as a prerequisite for students being able to comment intelligently on the subject domain. This generally involves the student's individual accumulation of teacher presented concepts. However, contemporary perspectives on learning suggest that learning is most effective when the learner is an active participant [12] [4], constructing content for themselves within a social system. Similarly, creativity within a domain involves the active construction of new ideas and content within the social context of other members of the field. Few scholars have suggested there exists an inherent relationship between learning and creativity. However, the similarities between the two are evidently striking [11] [18] [9]. It is suggested here that theories advocating learning as a socially constructive process may shed light upon creative phenomena. Extending up this, a distillation of creativity theory is presented in the form of an integrative framework. This framework exists as a design support tool in planning lesson materials for the classroom. This is demonstrated through SoundScape, a creative collaborative music composition program.

Constructivist learning theory, constructionism, creativity, collaboration, music learning

1. INTRODUCTION

Traditional pedagogy isolates the learner from social interaction and concerns pre-packaged lesson materials being delivered from the teacher and/or learning program to the student. Such an approach concerns itself with the passive absorption of knowledge, which is later tested via exam based scenarios. Although this may equip students to pass exams, they may face difficulty when applying concepts into authentic practice [2]. We therefore emphasise the importance of designing educational technologies in a way to facilitate the natural learning process.

2. THEORETICAL BACKGROUND: LEARNING AND CREATIVITY

With growing advancements in technology, learning programs are an ever present element of education today. However, technology is often misconstrued as a medium for disseminating knowledge to students as opposed to providing a virtual space in which the student is an active participant, exploring a domain for themselves. It is therefore emphasised that the focus of educational media should not reside with what technology will improve education, rather the way in which such technology is designed should be considered. This emphasises the importance of design considerations of e-learning systems. This paper extends upon constructivist and constructionist perspectives on learning. These perspectives suggest learning is not solely an individual process as people naturally interact with others and their surroundings, and learning is the outcome of these interactions [19]. Furthermore, from a constructionist point of view, it is important for students to be actively engaged in personally creating a product meaningful to themselves and others [14] [10].

2.1 The Creative Process

Wallas (1926) formalised the four stage model, representing the creative process [20]. This model consists of four stages; preparation, incubation, illumination and verification. Preparation concerns immersing one's self within a domain and developing a curiosity about a particular problem [7]. At this stage, an individual will also consciously accumulate knowledge and draw upon influences from previous experience. During the incubation stage, conscious thought pertaining to the problem is rested and left to the unconscious mind [3]. Illumination occurs when one experiences a sudden flash of insight [15]. Finally, verification concerns forming judgements pertaining to the creative artefact produced. A number of scholars have continued to apply the four stage model as a basis for understanding creativity [13] [8], while others have extended upon it [1] [16] offering models consisting of several stages. Others have proposed different approaches which show no correlation with the traditional model [5].

3. AN INTEGRATIVE FRAMEWORK FOR LEARNING AND CREATIVITY

Drawing on the above, a framework which represents a distillation of creativity theory is presented, focusing upon education. This framework is presented the form of an integrative framework, which exists as a design support tool to assist the design of creative educational experiences for the classroom (see figure 1). Wallas's four-stage model has been adapted as the fundamental basis for this framework, with the processes of preparation, generation and evaluation represented laterally across the framework. The vertical dimensions reflect individual (denoted here as personal) and social components of creativity. The 'social' level refers to others, peers and society. Whereas, 'personal' levels reflect explicit and tacit levels of thinking.

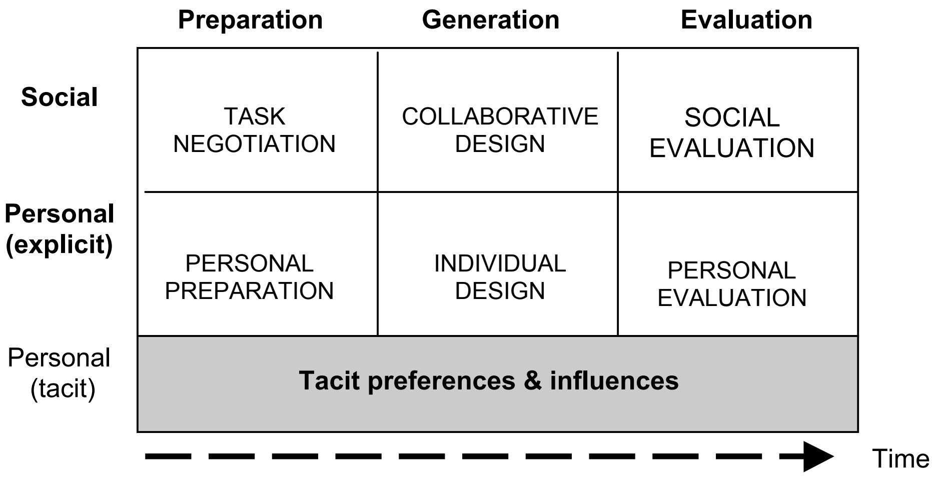


FIGURE 1: An integrative framework for learning and creativity

The processes of preparation, generation and evaluation are three integral concepts of the creative process. Every creative act involves the preparation of ideas. At a personal level, an individual will develop a curiosity or a desire to create. Once this desire has been established, information is consciously accumulated from the external environment and thoughts may be discussed with others on a 'social' level which the individual can reflect upon. If working in a collaborative setting, group-wide negotiations of the task will also take place. Inevitably, the way in which an individual prepares for the task will be influenced by their past experiences [17].

The generation process of the framework encompasses social and personal design. Within this process ideas are generated which can involve negotiation between the individual and peers in their environment. Additionally, idea generation is assisted partly by a continuous dialogue which occurs between conscious thought at the personal explicit level and sub-conscious processing at the tacit level. The evaluation process concerns reviewing early creative ideas through to evaluating the final artefact. Evaluation may occur at a personal level, or at a wider (community) level. It is emphasised that the framework does not commit to a strict linear route, rather the creative process is cyclic in nature. Therefore, the review of creative ideas may result in a need to revise ideas which may result in further preparation, or evaluation or further generation and so on. The processes of the framework are not mutually exclusive, as in some instances processes within the framework may overlap. The framework can be used as a design support tool to facilitate creative thinking in the classroom by ensuring that preparatory materials are scaffolded to the six component boxes of the framework.

4. APPLYING THE FRAMEWORK TO SOFTWARE DESIGN: SOUNDSCAPE

SoundScape is a music composition program which has been constructed as a vehicle to demonstrate how the framework can be applied in practice. It has been specifically designed for school aged children, allowing them to work collaboratively and creatively to construct a piece of music. SoundScape replaces traditional stave notation with 'themes' and 'objects'. Thus, constraints of musicality are removed. Students begin their interaction with SoundScape within the preparation process of the framework. Initially, students are set the task of selecting one of four themes including; a street, a jungle, an ocean and a space theme. Following their selection they are then presented with ten cartoon objects associated with the theme, which they must then match to music samples (see figure 2). Within this section of the 'preparation process', students can be expected to discuss the task to be completed within the paired-groupings. The program seeks to provide reflexivity in learning, by encouraging the students to think on a deeper level to justify their learning choices made. Therefore, when a picture and sound association has been made, the system will ask the students to explain why they have made that association. The composition interface is the point at which students enter the 'generation processes' of the framework. The interface relays to the student the selected theme which is set as the background and the selected objects are presented in coloured boxes at the bottom of the screen. The lines running from top to the bottom of the composition screen represent bar lines, so it is easy to depict images which are associated with a longer sound duration than others (see figure 3).

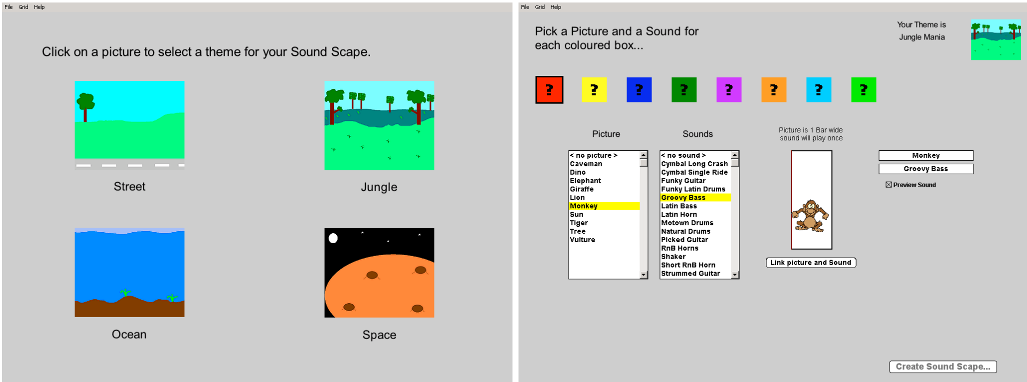


FIGURE 2: Software design: Designing to explicitly support preparation

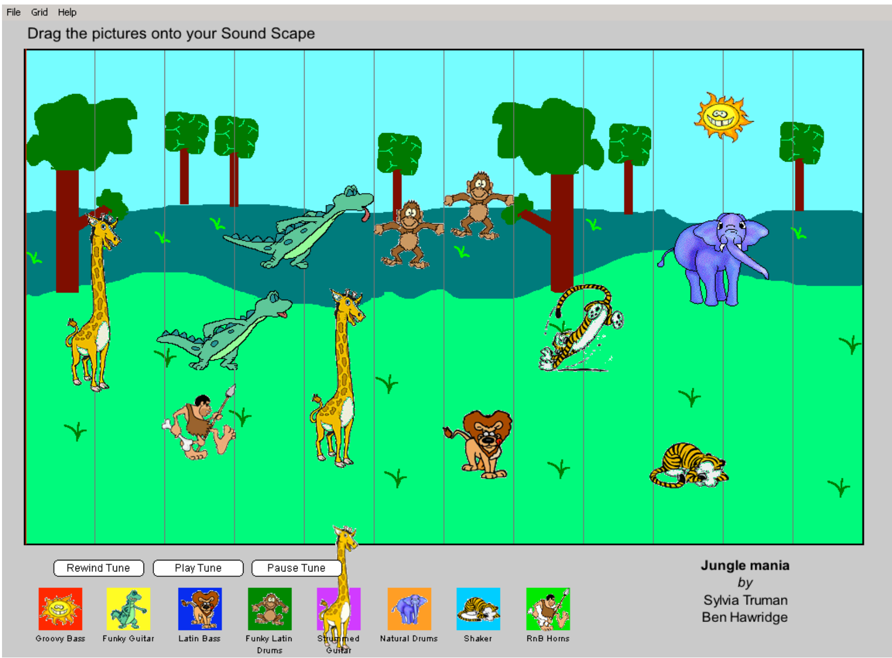


FIGURE 3: Software design: Designing to support generation

With regard to figure 3, students simply drag the objects from the coloured boxes onto the theme and structure them on the composition background as they wish. In terms of the framework, it is expected that students will collaboratively discuss and personally construct ideas. It is also expected that pair wise discussions may also trigger further realisation of ideas. In terms of evaluation, it is expected that on an individual level, a student will form their own judgements concerning the composed work. On a collaborative level, it is expected that pair-wise reflection and judgements concerning the composition will take place. Arising from this, students may move between generation and evaluation phases as refinements are made to the composition. Students might then seek wider evaluation of their composition from their peers and /or teacher. For example, students can listen to each others compositions or can print out the pictorial representations which can be exhibited in the classroom to encourage peer-wide evaluation.

5. CONCLUSIONS

The generative framework introduced and discussed here, is supported by Feldhusen & Treffinger's facilitators of a classroom environment conducive to learning [6] which include facilitators such as allowing students to explore concepts and learn from errors and allowing time for students to think about and develop their creative ideas. Additionally, it is important to encourage students to share, develop and learn together as well as independently as this allows students to make choices and have control of their learning. The above instantiation of the creativity framework considers these points. Thus, the generative framework of creativity can be applied to many differing domains of study within the classroom context, and many instantiations can be made.

This paper has outlined the motivation behind the framework and demonstrated how the framework can be applied in practice. The music composition program 'SoundScape' has also been presented as a vehicle to demonstrate how this framework can be applied in practice. The framework exists as a design support tool for educators, teachers and designer of educational technologies alike in the preparation and design of learning materials. The framework can be applied by mapping preparatory learning materials onto the six component boxes of the framework. This is part of a wide programme of research for which results are currently being analysed.

6. ACKNOWLEDGEMENTS

The authors acknowledge and extend appreciation to Ben Hawkridge (Knowledge Media Institute, The Open University) for his assistance with developing the SoundScape program.

REFERENCES

[1] Amabile, T.M (1996) *Creativity in Context.* Boulder, CO; Westview.

[2] Brown, Collins & Duguid (1989) Situated Cognition and the Culture of Learning. E*ducational Researcher.* January – February. 32-43

[3] Claxton (1998) *Hare Brain Tortoise Mind: Why Intelligence Increases When you Think Less*. London. Fourth Estate Limited.

[4] Duffy, T. M., & Cunningham, D. J. (1996). Constructivism: Implications for the design and delivery of instruction. In D. H. Jonassen (Ed.), *Educational communications and technology* (pp. 170-199). New York: Simon & Schuster Macmillan

[5] Eindhoven, J.E & Vinacke, W.E (1952) Creative processes in painting. *Journal of General Psychology.* 47. p 165 – 179.

[6] Feldhusen, J. F & Treffinger, D. J (1980) *Creative Thinking and Problem Solving in Gifted Education.* Dubuque, IA; Kendall/Hunt.

[7] Getzels, J.W (1964) Creative Thinking, Problem-Solving, and Instruction. In E.R. Hilgard (Ed), *Theories of Learning and Instruction*. Chicago; University of Chicago Press.

[8] Goswami, A (1996) Creativity and the quantum: A unified theory of creativity. *Creativity Research Journal.* 9. p 47 – 61.

[9] Guilford, J. P (1950) Creativity. *American Psychologist.* 5. pp 444 – 454.

[10] Harel, I (1991) *Children Designer's: Interdisciplinary Constructions for Learning and Knowing mathematics in a Computer Rich School.* Norwood; NJ. Ablex Publishing.

[11] Karnes, M. B, McCoy, G.F, Zehrbach, R.R, Wollersheim, J.P, Clarizio, H.F, Costin, L & Stanley, L.S (1961) *Factors Associated with Overachievement of Intellectually Gifted Children.* Champaign, IL.

[12] Lave, J & Wenger, E (1991) *Situated learning: Legitimate Peripheral Participation*. Cambridge; Cambridge University Press.

[13] Osche, R (1990) *Before the Gates of Excellence: The Determinants of Creative Genius.* New York; Cambridge University Press.

[14] Papert, S (1993) *The Children's Machine: Rethinking School in the Age of the Computer*. New York. Basic Books.

[15] Poincare (1913) in Leytham, G (1990*) Managing Creativity.* Norfolk. Peter Francis Publishers.

[16] Runco, M.A & Dow, G (1999) Problem Finding. In Runco, M.A & Pritzker, S.R (Eds) *Encyclopedia of Creativity.* (Vol. 2). P 433 – 435. San Diego, CA; Academic Press.

[17] Schank, R (1995) *What we Learn when we Learn by Doing.* Northwestern University, Institute for Learning sciences.

[18] Torrance, E.P (1981) Creative teaching makes a difference. In Gowan, J.C, Khatena, J & Torrance, E.P (Eds) *Creativity: Its Educational Implications.* Dubuque, IA; Kendall/Hunt

[19] Vosniadou, S (1996) Towards a Revised Cognitive Psychology for New Advances in Learning and Instruction. *Learning and Instruction.* Vol. 6. No. 2. 95 – 109

[20] Wallas (1926) *The Art of Thought.* London; Jonathan Cape [republished in 1931].

Interactive Experiences

Authentication Using Tactile Feedback

Ravi Kuber and Wai Yu
Queen's University Belfast, Belfast, BT7 1NN, UK
{r.kuber, w.yu}@qub.ac.uk

As current knowledge-based systems do not take into account human limitations when recalling authenticating strings, individuals often select short or guessable passwords. As a result, a compromise is created between memorability and security. To address these issues, a novel approach has been developed examining the use of recognition-based authentication, through the use of the haptic modality. The Tactile Authentication System (patent pending) allows the user to authenticate entry through the ability to recognize previously perceived tactile sensations. This method of authentication presents benefits over alternative visual-based systems. Stimuli are perceived underneath the fingertips, protecting the user from the threat of observers recreating the authentication sequence. As the sense of touch is personal to each individual, tactile sensations are difficult to describe in concrete terms, so cannot be easily shared with others. This paper reports the design of the Tactile Authentication System (TAS), and presents results from a one-month trial. Participants were able to recognise their pre-selected tactile stimuli from a wider range of sensations, and could form their personal sequences to authenticate entry to the system. TAS has been shown to present a memorable and usable alternative to conventional knowledge-based authentication systems.

Authentication, memorability, haptics, perception, tactile

1. INTRODUCTION

The need for heightened security is intensifying as organisations begin to automate more business functions, which in turn, increases the amount of sensitive data stored in electronic format on shared networks. Renewed security threats and legislation such as the Data Protection Act (1998), have driven organisations to realise the importance of information security. Knowledge-based authentication (KBA) is the most common method of user authentication, where the trusted individual identifies himself/herself with a username, and authenticates entry using an alphanumeric password [4]. Identification establishes the user's right to access the system. The authentication mechanism verifies that the user is the legitimate owner of the ID [2]. KBA is favoured over alternatives such as biometric authentication, as it is a well tested form of technology, simple to administer, well understood by users and system administrators, and requires no additional hardware or software [11].

Whereas in the past, organisations have mainly concentrated on securing data from the threat of physical attacks, the focus has shifted to issues associated with the user (human factors). This is due in part to limitations of the human memory. With the proliferation of technology, individuals need to increasingly authenticate access to multiple systems on a daily basis [2]. Precise recall of information is known not to be a strong point of human cognition [3]. Natural decay of information from our memories and within-list interference effects from other similar chunks of information can occur, meaning that the user may find it difficult to authenticate entry into systems, particularly when recalling randomly generated passwords. As a result, individuals often tend to select 'weaker' passwords, which are often short and guessable. In an attempt to achieve memorability, security is compromised. The threat of third-party attacks with KBA mechanisms is also evident. Individuals have been found to write-down or share authentication information with others [1]. 'Shoulder-surfers' monitor the spatial position of keystrokes made by users on numerical-keypads and keyboards. In both cases, authentication sequences can be recreated and system entry can be achieved. A need has been identified for a system addressing issues of weaknesses arising from human factors and observer attacks.

2. GRAPHICAL AUTHENTICATION & THE TACTILE MODALITY

Studies have shown that pictures can be easily committed to memory [2,3,11]. In contrast to recall-based systems used for remembering alphanumeric passwords, systems employing graphical authentication make use of the benefits offered by recognition. Recognition involves a less resource-intensive process when compared to recall, so would benefit the end-user when interacting with authentication technology. The Déjà Vu system [3] has been designed to effectively recognize abstract art images in a sequence (portfolio), from a larger set of images presented by a server. The Passface system [2] works in a similar fashion, asking users to remember and cognize photographs, exploiting our exceptional abilities to recognize faces. Whereas a recognition-based approach can lead to a lower level of mental workload being expended, scanning through each item visually can prove to be a time consuming process. This possibly accounts for the limited market share of graphical-based authentication mechanisms, compared to alphanumeric-based ones.

Research has shown the human ability to remember haptic information [10], however relatively little is known about the sensory memory store responsible for retaining haptic data [5]. Estimates dictate that our tactile memory span is between two to three items [14]. Mahrer & Miles [8] found participants were able to recall between four to six tactile stimuli, presented through the form of taps to the fingers. The researchers have also discussed the benefits that vision can bring to the tactile recognition process [9]. Relatively little research has been conducted in the areas of tactile recognition and the long term capacity of tactile memory. However, it is thought that the tactile channel could provide benefit for purposes of authentication.

3. SYSTEM REQUIREMENTS

The Tactile Authentication System (TAS) has been developed in order to provide an alternative to alphanumeric and graphical authentication systems. It makes use of our tactile memory capabilities to remember and recognize pre-selected tactile stimuli. The system has been designed according to the following criteria, discussed in greater detail by [1, 2, 3]:

- **Memorability**: The system should benefit from the strengths offered by recognition, compared with recall.
- **Security**: Users should not be able to select ‘weak’ authentication information. Design should focus on weaknesses arising from human factors. Threats from observer attacks should be minimised, and authentication information should be difficult to externalise.
- **Usability**: The system should be easy to learn, providing the user with a usable and stimulating experience.

4. SYSTEM DESIGN

The VT Player device (Figure 1) has been chosen for the purposes of the demonstration, as it works in a similar manner to a computer mouse, and has the benefit of providing customised tactile output through a pair of electronically-controlled matrix units (contactors) built into the surface just where the fingers rest. The matrices roughly resemble the display of Braille cells. Each 4x4 matrix holds 16 pins which rise and fall dynamically, gently delivering a tactile sense of the screen to the user’s fingertips. Pins can be arranged to form static patterns (Figure 2), or alternatively can be designed to change state over time, providing animated-style sensations. The user places his/her index and middle finger over the contactor pads to perceive the tactile stimuli. As participants’ fingers cover the entire contactor pads, patterns are visually-occluded when using the device.

The Tactile Authentication System [6] interface has been developed using the following Web technologies; HTML, Javascript, VB Script, and the VT Player SDK, enabling users to authenticate access to a system on a standalone PC. The Web application environment has been simulated using Microsoft Internet Information Server 5.0, allowing the standalone PC to act as a server. The tactile sensations used in TAS have been developed in view of the findings from an earlier perceptual study conducted by [7]. Distinctive static and animated pin patterns have been designed to enable users to quickly and accurately denote differences between each form of stimuli. Pin stimuli in the form of visual objects such as geometric shapes, lines, and symmetrical patterns have been included, as users were able to recognise these effectively in [7] (Figure 2). Pins were not positioned in close spatial or temporal proximity, in order to reduce effects of tactile occlusion.

FIGURE 1: VT Player device

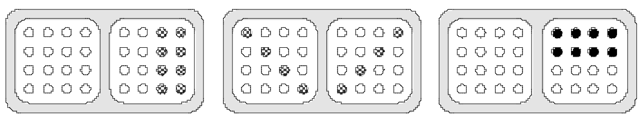


FIGURE 2: Examples of pin stimuli used in experiment

5. ENROLLMENT & MAIN TRAINING

All participants are provided with initial training on the VT Player device to expose them to a wide array of tactile stimuli and test basic recognition abilities. They are asked to select four tactile stimuli, from a selection of thirty-six distinctive sensations. Participants are then questioned on their reasons for selection, and given the opportunity to feel and memorise the tactile stimuli, in sequence format. They are then presented with a series of stimuli, including their own (main training phase), and are asked to identify their personal tactile sensations, in the sequence originally selected. Participants are then asked to repeat this procedure ten times, to commit the authentication sequence to memory.

6. AUTHENTICATION TO TAS

The authentication process employed by TAS is based on the method used by [2]. In order to successfully authenticate entry into the system, the participants select their name from a drop-down list, and then choose their personal tactile authentication sequence, comprising of four tactile stimuli, in the order originally selected in the enrollment stage.

Participants are presented with four grids, each of which contains nine visually-identical squares arranged in a 3x3 format (Figure 3). Each square contains a static or animated stimulus from the original thirty-six pin patterns designed for the study. As the users hover over each square, a different stimulus is perceived underneath the fingertips. Participants are asked to explore the sensations present in the first grid, and select one stimulus which corresponds to a sensation in their own personal authentication sequence. The process is repeated on the remaining three grids, until four stimuli have been selected (Figure 4). The order of grid presentation remains constant, as does the nine tactile stimuli contained within each grid. To minimise the chance of participants memorising the visual position of squares containing tactile stimuli, the order of sensations presented within a grid is randomized. Using four tactile stimuli selected from four grids provides 9^4 combinations, meaning that the chance of someone guessing a sequence at random would be 1 chance in 6561.

The user can try up to four times to enter the authentication sequence, before a sequence reminder containing four tactile stimuli, is sent to the participant. This is similar to the system employed by banks, where three PIN attempts can be made using the same card, in any one ATM machine. Upon receiving a reminder, the participant would be asked to go through the training procedure again, to commit the sequence to memory.

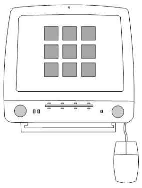


FIGURE 3: TAS system displaying grid of tactile stimuli

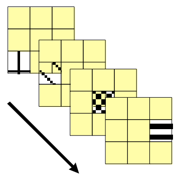


FIGURE 4: Example of authentication sequence to enter system

7. EXPERIMENT DESIGN, FIELD TRIAL RESULTS & DISCUSSION

A field trial was conducted with sixteen participants (12 male, 4 female, aged between 23 and 42), ten of whom had accessed the VT Player device one month prior to the trial. The trial aimed to assess the following:

- Recognition of tactile stimuli sequences over a short and long term period.
- Usability of the TAS interface and general system.

The participants were asked to log-in to the system every working day (Monday to Friday) for two weeks, and once at the end of the fourth week, following a procedure adapted from [13, 2]. Both the enrollment and authentication stages were completed in secure environments away from third-parties, as specified by [3].

Participants took up to four minutes carefully selecting tactile stimuli, in an attempt to build-up an authentication sequence. The reason behind the lengthy enrolment time was due to the fact that participants wanted to firstly perceive all tactile stimuli available, and make an informed choice as to which they thought the most distinctive sensations were to be. It was also evident that participants selected static, or a mixture of static and animated tactile stimuli, to form their authentication sequences. When questioned on their choice of stimuli, most replied that distinguishing between animated stimuli was a tougher process, compared with differentiating between static feedback. This could have been attributed to the participants' unfamiliarity with haptic devices and temporal limitations which made distinguishing between animated pin stimuli a tougher process for the users. Further analysis of authentication sensations used, revealed that participants were not selecting the same pin patterns as each other, opting for diverse 'static' or 'static and animated' choices.

	% remembered (initial attempt)	% remembered (successive attempts)
Day 1	100% (1st attempt)	N/A
Day 2	93.8% (1st attempt)	100% (4th attempt)
Day 3	78.6% (1st attempt)	100% (2nd attempt)
Day 4	91.7% (1st attempt)	100% (2nd attempt)
Day 5	100% (1st attempt)	N/A
Day 8	87.5% (1st attempt)	100% (2nd attempt)
Day 9	92.3% (1st attempt)	100% (2nd attempt)
Day 10	100% (1st attempt)	N/A
Day 11	100% (1st attempt)	N/A
Day 12	92.9% (1st attempt)	100% (2nd attempt)
Day 28	84.6% (1st attempt)	100% (2nd attempt)

TABLE 1: Successful attempts to authenticate entry

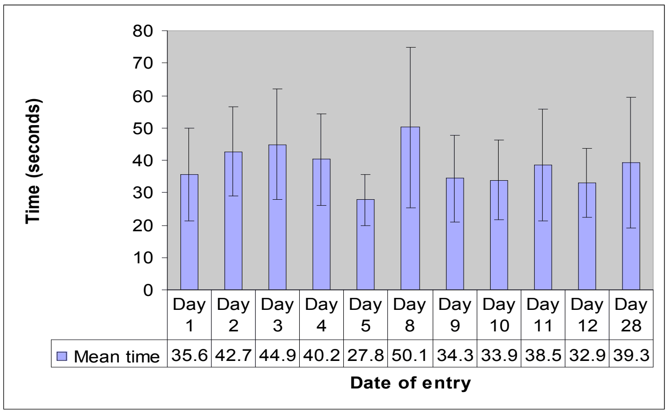


FIGURE 5: Time taken to authenticate access using tactile feedback

After selecting their stimuli, users were asked to enter their chosen tactile sensations, ten times in sequence order. The system would help the users to recover from any errors made during these practice sessions, by visually highlighting incorrect selections and directing the participants towards the square containing the original stimulus. By the fourth attempt, participants managed to authenticate their entry into the system free of errors, indicating a priming stage may have taken place. Human implicit memory, indicated by priming, refers to a change in performance of some task, due to prior exposure to the task materials [5]. The implicit memory has been shown to extend to the haptic memory [12], indicating that our tactile abilities may have played a part in the encoding of tactile authentication sequences in the TAS study.

After the training period, all sixteen participants were successful at authenticating entry to the system within the first attempt (100% - Day 1). This was performed in an average entry time of 35.6 seconds (Figure 5). For the rest of the week, participants continued to attempt to enter the system, with low levels of error (Table 1). Users could work their way through the nine stimuli contained within a grid, quickly enough to perceive the pattern presented within a short period of time and make a decision whether to select, or to move to the next stimulus. Interestingly, after gaps without rehearsal of authentication sequences, participants were still able to accurately authenticate entry to the system within the first few attempts, highlighting the long-term memory aspects of non-textual feedback, also found in studies by [13]. It was noted that participants would spend longer after a gap of the weekend (50.1 seconds – Day 8), carefully selecting stimuli from the choices available. However, the time taken to authenticate entry would reduce on successive days. By the end of the month, with over two weeks without using the tactile stimuli, all participants were able to login within the second attempt. This shows that the tactile channel can be beneficial in both the short and long term for purposes of authentication.

Participants were generally found to be successful at authenticating entry on the first attempt (92.9%). Throughout the course of the one month trial, users were found to make 13 incorrect attempts to access the system from a total of 153 recorded entries (8.5%). 7 self-aborted attempts were also recorded. Further analysis revealed that a small number of participants were becoming confused between two similar stimuli, which they could not distinguish between. This was most likely due to the effects of occlusion. Brostoff and Sasse [2] have noted that failed login attempts are user costs, and so should be minimized where possible. As all participants were able to log-in within the four attempt limit, no reminders needed to be sent out.

Regarding the question of how were users able to remember their stimuli, many participants admitted that as they were able to explain their choice of stimuli verbally, and make a visualization of the patterns in their minds. This may have aided them to retain authentication information over the month-long trial period. Mahrer and Miles [9] have suggested that memory for a sequence of tactile stimuli involves the deployment of strategies using a combination of verbal rehearsal and visio-spatial recoding rather than relying solely on the retention of tactile sensations. A similar situation may have also applied to TAS users.

In terms of usability, TAS was rated quite favourably as a method for authenticating entry. The system was found to be understandable, learnable, and operable without the need for technical support. However, participants did point out that isolating each stimulus within each grid was a time-consuming process. An average of 38.2 seconds was taken to authenticate entry to TAS. Using a recall-based approach would be considerably faster. It is thought that by providing salient feedback which can be perceived instantly on a less visual-centric interface, may bridge the gap between time taken to enter both types of mechanism. Participants were generally pleased with the strengths that the randomized presentation of information on the Tactile Authentication System interface brought. It led them to feel more secure from the threat of observer attacks. Tactile feedback experienced underneath the fingertips was perceived as a secure method of presenting information away from shoulder-surfers. With the recent switch-over to Chip'n'Pin technology to the UK and the growth of ATM machines, users saw promise in using tactile feedback to occlude visual information from onlookers when making financial transactions. Tactile authentication technology was thought to provide an inclusive authentication solution, allowing the visually-impaired community to access private data more effectively. Whereas participants of the TAS study highlighted the benefits and feasibility that tactile authentication has to offer, the tactile authentication trial was not completely representative of the memory demands placed upon users. This was the first tactile authentication sequence that participants were asked to remember, so there were no issues from within-list interference. Future work will aim to assess how effectively multiple authentication strings can be recognised, and whether interference effects will occur in these situations.

8. CONCLUSIONS & FURTHER WORK

The paper has described the design and development of a novel authentication system that makes use of our ability to memorise and recognise tactile feedback. Results from a study have shown that sixteen participants were able to memorise and authenticate entry for a month long period, with low levels of error. The sense of touch is unique to each user, making tactile stimuli more difficult to write-down or disclose to others. Tactile stimuli are occluded from view when entering the system, so it will be tough for third parties to observe the sensations. The next logical step would be to evaluate TAS over a longer period of time, to obtain a greater representation of the capabilities of the human memory, when using tactile feedback. The findings can then be compared with those of alphanumeric and graphical authentication. It is thought that the human-computer interaction of TAS could be enhanced with a more suitable interface design, and by providing more recognizable forms of tactile feedback on larger areas of skin. In this way, the tactile stimuli could also be perceived within a shorter time frame.

REFERENCES

[1] Adams, A. & Sasse, M. A. (1999) Users Are Not the Enemy. Communications of the ACM 42, 12, pp. 40-46.

[2] Brostoff, S. & Sasse, M.A. (2000) Are passfaces more usable than passwords? In Proc. HCI'00, pp. 405-424.

[3] Dhamija, R. & Perrig, A. (2000) Deja Vu: A User Study Using Images for Authentication. In Proc. USENIX'00.

[4] Garfinkel, S. & Spafford, G. (1996) Practical Unix and Internet Security, O'Reilly & Associates, USA.

[5] Klatzky, R. L. & Lederman, S. J. (2002) Touch. In Experimental Psychology, Wiley, New York, USA, pp. 23-25.

[6] Kuber, R. & Yu, W. (2006) Tactile Authentication. Patent No 0603581.0 (Patent Applied For).

[7] Kuber, R. & Yu, W. (2006) Using Tactile Feedback for Authentication (In process of submission to IJHCS).

[8] Mahrer, P. & Miles, C. (1999) Memorial and Strategic Determinants of Tactile Recency. Exp Psy, 25, 3, 630-643

[9] Mahrer, P. & Miles, C. (2002) Recognition memory for tactile sequences. Psychology Press 10, 1, 7-20.

[10] Millar, S. (1999) Memory in touch. Psicothema, 11, 747-767.

[11] Renaud, K. & De Angeli, A. (2004) My password is here! Interacting With Computers, 16, 6, 1017-1041.

[12] Srinivas, K., Greene, A.J. & Easton, R.D. (1997) Implicit and explicit memory for haptically experienced two-dimensional patterns. Psychological Science 8, 243-246.

[13] Valentine, T. (1998) An Evaluation of the Passface Personal Authentication System (Tech Report). London.

[14] Watkins, M.J. & Watkins, O.C. (1974) A tactile suffix effect. Memory & Cognition 5, 529-534.

Andrew Rivolski: Cooperative Multi-screen Network Game

Andy Yamada, Tomoyuki Nezu and Masa Inakage
Keio University Inakage Laboratory, 5322Endo Fujisawa Kanagawa japan 252-8520
Phone/FAX +81-466-49-3545
andy@imgl.sfc.keio.ac.jp, tomoyuki@dlb.sfc.keio.ac.jp, inakage@sfc.keio.ac.jp

Andrew Rivolski is a multiplayer network game played in an environment consisting of multiple displays. The players are placed in two different remote locations, and by having mutual interactions the game allows the players to experience a cooperative phenomenon.

Leisure and games, collaborative working, interaction techniques

1. INTRODUCTION

Imagine two groups of people placed in two different locations with no means of contact with each other but just a single location signal of each other shown on an immersive panoramic image projected on to 10 screens surrounding them. As long as these two groups of people have a common goal and reaches towards this goal together by thinking and moving as if they were one another, a cooperative phenomenon will occur and unite the remotely separated group of people through a new form of communication. By using modern technology such as the Internet, Andrew Rivolski, a multiple player cooperative network game will provide a completely innovative form of communication to game players.

1.1 Objective of the game

The objective of our multiple player cooperative network game is for the players located in different locations to socially interact with each other and foster the relationship between each other. As communication by telephone and E-mail become more popular and wide spread, we tend to forget how important a simple act such as a handshake or a hand wave good bye can be in communication. Instead of looking one in the eye and apologizing, today, we can just type a simple “sorry” note by E-mail and one’s bad act of conduct can be dismissed. One can just call in “late”, 10 minutes before a meeting and one’s effort of getting somewhere on time and not giving one discomfort can be dismissed. Because we take advantage of the convenience that modern technology provides us, not only are people forgetting the importance of mutual greetings like a wave good-bye, respect for others is beginning to lack. True communication should be presented directly and not through technology such as a mobile phone or E-mail. Small mutual interactions like these are what make the communication real and yet because of the evolution and rapid diffusion of technology, this is what we tend to lack. By adopting these mutual interactions as a rule for the game, and offering the players the mentality of showing respect to the other players, our contents will be a resolution to the problem and bring an affluent daily life back to the people.

2. GAME DESCRIPTION

The players are separated in to two remotely distant rooms and are not able to see each other. Players of 2 to 4 are placed in each of these rooms and the players must cooperate with one another to reach a common goal. Actual movements of the players reflect on to the interface and the corresponding team understands how the other team is thinking and wanting to move. Three of the walls of the room will function as displays. (Figure 1.) The displays consist of 10 individual displays connected together to make one long panoramic image. (Figure 2.) Figure 2 is an example of the 10screen version of the panoramic game image and Figure 1 is an example of the 4screen version of the panoramic game image. From these displays, the two teams of players will view the same interface, meteors flying from side to side towards a lonely satellite placed in the center. The players of each location must cooperate with each other by using the 9 floor panels placed in each other's rooms to maneuver the satellite with out being hit by meteors. Each floor panel is 91cm x 91cm and placed on the floor so that there are 3 rows and 3 columns of the panels like Figure 1. The floor panels are this size because it is just about the perfect size for 4 players to mush up in. By giving hints of which way each other are going to move to each other, the players will inevitably begin to move in the same direction as each other and when they move in the same direction at the same exact moment, a cooperative phenomenon will occur and the players will be rewarded in the game.

FIGURE 1: Andrew Rivolski (4 Screen Version)

FIGURE 2: Andrew Rivolski (10 Screen Version)

3. USER ACTIONS

When all the players in one of the rooms move to a certain panel, a signal will be sent to the satellite inside the game. When the players in the corresponding room all move to a certain panel, another signal is sent to the satellite. Then the game will take the sum of the two team's signal and move in the necessary direction. If one team's signal was to move 2 steps to the left and the other team's was 3 steps to the right, the satellite will take the sum of that and move 1 step to the right. Users must note that in order to move the satellite in the intended direction, they must predict the intended direction of the other team of players.

The panoramic interface consisting of 10 screens will add to the challenge of the players. 3 screens each on the side and 4 screens in front of the players will create an environment in which the players must keep an eye on all the screens surrounding them. If all the players are caught up in trying to move the satellite away from the closest meteor that is heading towards the satellite, they might move right in to the path of another meteor that had entered the side screens. Team work of watching out for meteors from all over and quick decision making on which panel to move to will be needed.

Quick decision making between the players will be necessary, not only do the players have to think and act so that the satellite will move in the same direction both teams want, the players within each of the rooms must decide on which panel to move to. The signal will only be sent to the satellite when all players are standing on one panel. If the players cannot come up with a decision quick enough, the satellite will most likely not move in the direction the other team has intended or their own team's intensions and will be attacked by the meteors. Hence, not only is cooperation needed between the corresponding teams, cooperation between the players in each room is also needed in order to maneuver the satellite in the direction both teams intend.

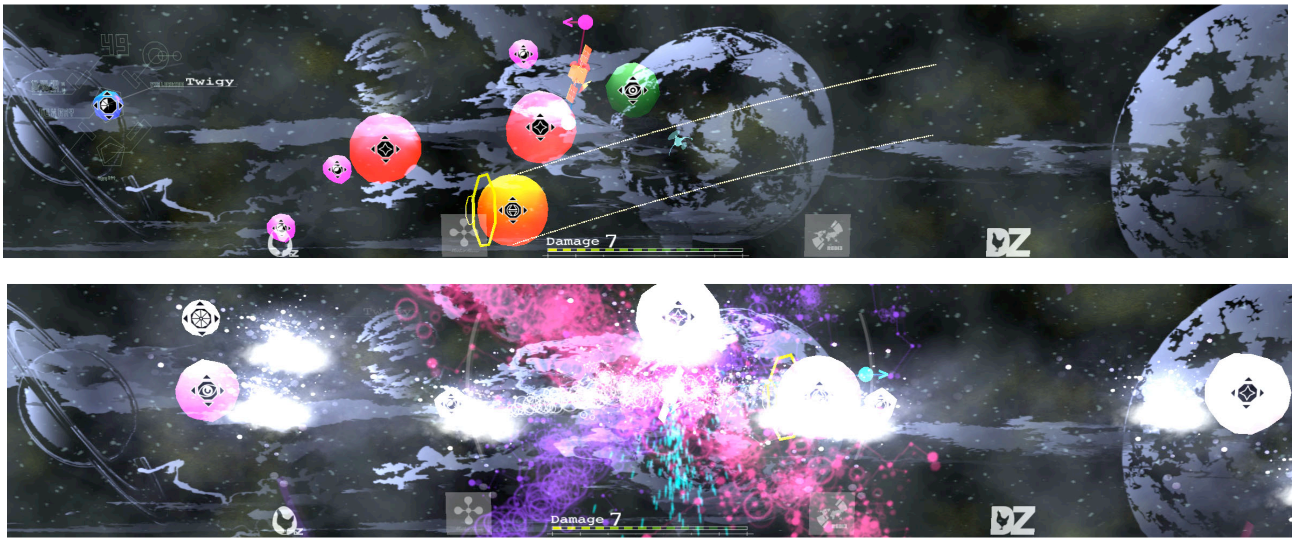


FIGURE 3: Immersive Game Interface using 10 screens

4. SYSTEM DESCRIPTION

For each panel that the players use as a controller, pressure sensors are used. The inputs sent from these 90cm x 90cm pressure sensors go through the iCube Midi digitizer and is processed through the software called “VVVV”. “VVVV” is a graphical programming free-ware for real time video synthesis. The graphics are rendered on to screen_using DirectX. At each location, one main machine is serving as a server, and 10 other clustered machines are used to render the graphics as a panoramic immersive environment, synchronized via network. The two main servers are connected via high speed Internet, to share and synchronize the two locations and transferring the live video feed in HD.

5. COOPERATIVE PHENOMENON

As the players play through each of the stages of the games, they will begin to get the feel of how the other team moves, and both teams will begin to move in the same direction as the other team to successfully maneuver the satellite. When the same signals from both of the teams are sent to the satellite at the same time, the cooperative phenomenon will occur. We believe the fact that multiple groups of people trying to reach a common, goal coming up and moving in a certain way at the exact same moment when they can’t see the other group of people is a phenomenon suggesting a new way of communication using modern technology. When this phenomenon occurs, a live video feed of the other team will be projected on to the screen in front of them and both teams of players will be able to see their corresponding team players for the first time. Receiving this live visual greeting from the corresponding team after thinking and playing so that both teams can reach the same goal will surely be a spectacular experience.

6. ACKNOWLEDGMENTS

This project is conducted at Research Institute for Digital Media and Content.

REFERENCES

[1] Mike Brinker, Jenova Chen, Vincent Diamante, Todd Furmanski, Erik Nelson, Glenn Song, Dyadin, DiGRA 2005 http://www.dyadin.com/
[2] Adrian David Cheok, Kok Hwee Goh, Wei Liu, Farzam Farbiz, Sze Lee Teo, Hui Siang Teo, Shang Ping Lee, Yu Li, Siew Wan Fong, Xubo Yang Human Pacman, *Proc Personal and Ubiquitous Computing 2004*, ACM Press(2004)
[3] Usman Haque, Sky Ear http://www.haque.co.uk/skyear/
[4] The SINE WAVE ORCHESTRA http://swo.jp/
[5] Yohoho! Puzzle Pirates, Three Rings Design,Inc http://www.puzzlepirates.com/

Evaluation of a Crisis Management Head Mounted Display (HMD) System

Huseyin Dogan, Ben Dawson, Elizabeth Carver, Jeremy Hinton
BAE Systems, Advanced Technology Centre, Human Factors Department, Sowerby Building
FPC 267, PO Box 5, Filton, Bristol, BS34 7QW, United Kingdom
Huseyin.dogan@baesystems.com

The aim of this research is to investigate the potential value of introducing Head Mounted Display (HMD) technology into real world crisis management teams. Helmet mounted display is particularly suitable for people that require instant and easily accessible information during their fieldwork. We interpret this as exploiting HMD capability as a component of a wearable information system to support decision making and to enable a field operator to share and receive information during a crisis event.

The construction and architecture of the functional HMD concept demonstration system is described. In addition, a qualitative evaluation of the various human factors and socio-technical issues that emerged when developing and trialling the HMD system is included in this study. The investigation comprised an exploratory experimental evaluation of the HMD system using human factors metrics and techniques such as CARS (Crew Awareness Rating Scale) and Workload measurement questionnaires.

crisis management, human factors, helmet mounted display, e-briefing, situational awareness, workload

1. INTRODUCTION

1.1. Project aim

This research is part of a European integrated project called OASIS (Open Advanced Systems for Crisis Management). (http://www.oasis-fp6.org/).

The aim of OASIS is to define and develop a first version of an open, modular and generic Disaster and Emergency Management (DEM) System in order to improve the effectiveness and efficiency of all agencies within the European Union who were likely to be involved in the management of Disaster and Emergency Operations (DEO). The aim is to specify and design a true generic, interoperable and open architecture which will allow easy deployment at every level of the action chain (local, regional, national and European). This generic architecture will rely on the integration of mature state-of-the-art technologies.

1.2. Subproject aim: e-briefing and tasking with HMD technologies

The aim of this research is to investigate the potential value of introducing Head Mounted Display (HMD) technology into crisis management teams. Helmet mounted display tool is for people that require access to information during their fieldwork. We interpret this as exploiting HMD capability *as a component* of a wearable information system to support decision making and to enable a field operator to share and receive information relating to a crisis.

A head mounted display (HMD) was demonstrated exchanging text, voice, video and graphics, and using voice control over a wireless connection via Netmeeting. The objective was to look at new technologies to:

- Support first responders/blue light forces.
- Communicate with the higher command centres and with their colleagues both in the same force as well as cross agency.
- Exchange information in a timely way to enhance safety and effectiveness.

This research comprised an exploratory experimental evaluation of the HMD system using standardised human factors techniques such as CARS (Crew Awareness Rating Scale) and workload measurement questionnaires.

2. THE SYSTEM

In this section we describe how the crisis management HMD system was constructed, the materials and technologies used and discuss the system architecture that was developed. A field operator was equipped with an HMD, camera, and communications link over a wireless Netmeeting application (see figure 1 on the following page). The system was designed in order to allow a wide range of data such as instructions, status information and guidance, to be transmitted using voice, text, graphics and video streams. Research suggests (Schneiderman, 1996; Carr, 1999) that supporting visualisation techniques can enhance levels of situational awareness as well as lessen the cognitive workload placed upon personnel. Also, voice commands have been developed to support a field operator's interaction with the system. Herdman et al (2001) address the cost benefits of the integration of DVI (Direct Voice Input) in complex systems. Voice commands proved to be successful in this trial, but there are still operational difficulties such as specific user training, background noise and command word confusability.

The Local HQ computer in figure 1 contains four different screen sections: video stream, text instructions, site map, floor plans and route directions:

- The video section can be a live feed received from the field operator's webcam or from local CCTVs.
- The text section contains written instructions which can be sent between the field operator and the HQ.
- The third section of the screen incorporates the site map which was designed to include information about the field operator's surrounding area and any local vehicles which were fitted with GPS tracking devices.
- Finally, the directions section of the HQ (Headquarters) screen was designed to contain images of floor plans and route instructions supporting the field operator to important locations (for example where there were reports of casualties). The field operator wearing the HMD can maximise and switch between any of these screens using voice commands.

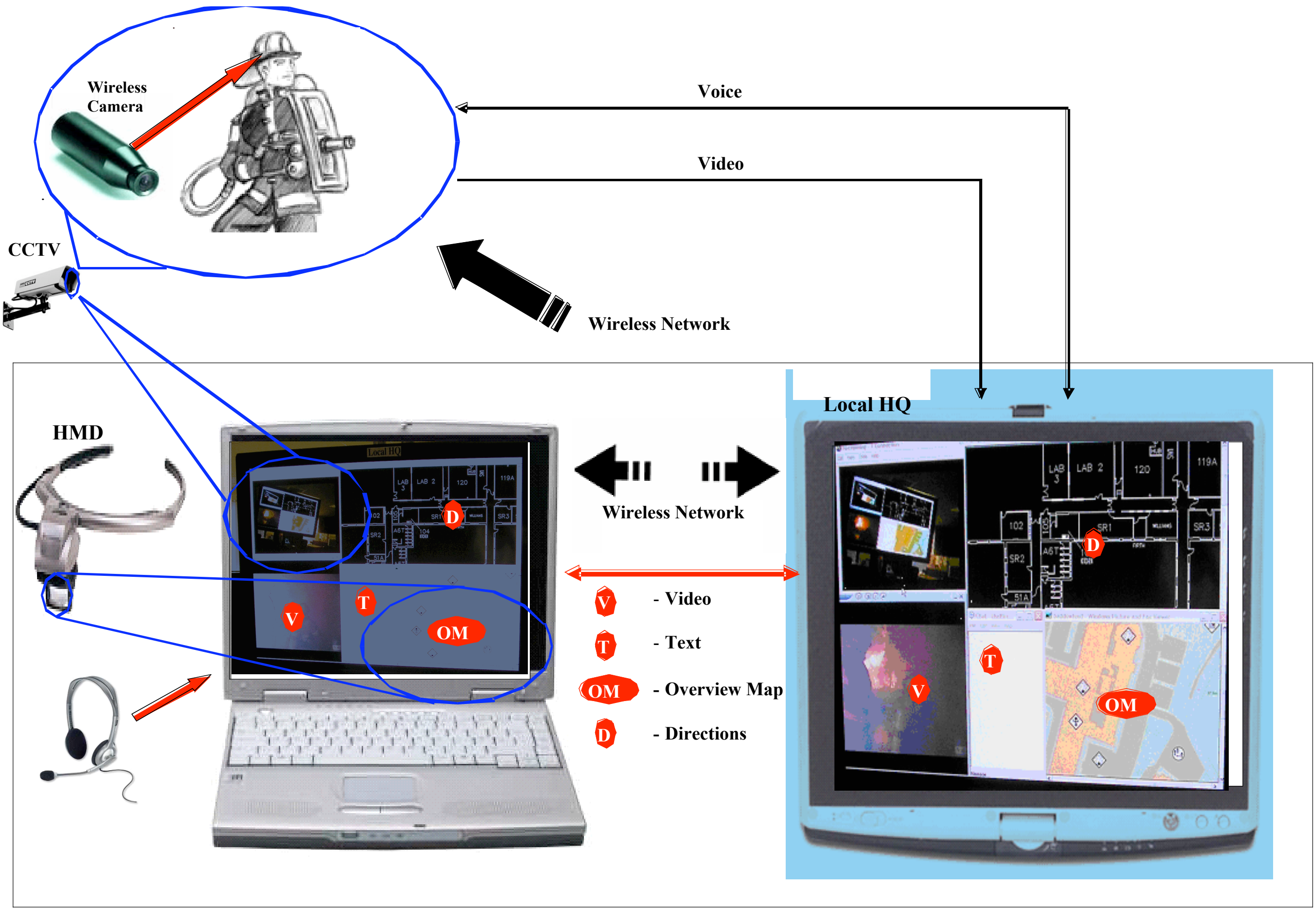


FIGURE 1: HMD and Local HQ displays with communication media

3. METHODOLOGY

FIGURE 2: Field operator and HMD

The HMD system (Figure 2) was used as a component of a wearable information system with the aim of supporting decision making, increasing situational awareness and reducing cognitive workload by enabling field operators and HQ staff to share and communicate critical information relating to crises. A pilot trial was carried out in order to assess the benefits of presenting information in this way. Larger user trials are planned but the preliminary trial helped identify where improvements needed to be made, and enabled the performance metrics to be trialed.

Consequently, it must be stated the experiment was a simple exploratory evaluation designed to establish whether the HMD system had the potential to be valuable as a crisis management support tool. We did not seek to determine statistical significance or attempt to prove whether use of the system had a significant impact on measures of situational awareness or workload. Instead, the aim of this experiment was to gauge whether there were any patterns to the results which suggest that further research in this area is warranted and whether effort in ironing out technical problems is worthwhile in future work.

3.1. Method

Nine participants were allocated to three conditions. In each condition participants were given the same instructions: they were told that they were participating in a crisis management exercise and that they were to play the role of a field operator. None of the participants were familiar with the layout of the building.

3.2. Participants and Procedure

9 participants were given a task of locating a casualty trapped in a building. Participants were supported by different media in each condition. The three different conditions were as follows:

- *Condition 1 – With everything:* HMD, video output, speech, directions (floor plan), overview map and text instructions. The directions section of the HMD contained ordered slides showing extractions from the floor plans. The field operator could navigate between these slides at any time using voice commands.
- *Condition 2 – Without the floor plan:* HMD, video output, speech, overview map and text instructions but *no floor plan* was given to the participants.
- *Condition 3 – Only given a paper copy of the floor plan:* the floor plans used were the same as used for Condition 1.

In addition to capturing time to complete the task, CARS (Crew Awareness Rating Scale) and workload measurement questionnaires were given to the participants during the exercise to measure how they were coping with the crisis. Typical questions included:

- *Would you say that you observed all events and information that were relevant to managing your task?*
- *Would you say that you had a good sense of the future course of events and likely outcomes with regard to completing your task?*
- *Would you say that you were aware of the best course of action?*
- *Would you say that you found it easy to decide exactly what to do?*

In summary, the following measures were taken:

- Time and time pressure
- Situational awareness
- Task completion
- Problems
- Confidence
- Mental Demand
- Effort
- Frustration

A member of the experimental evaluation team followed the field operator (participant) to observe and analyse the technical, behavioural and operational implications of using the OASIS HMD system.

4. RESULTS AND DISCUSSION

Differences between groups should not be taken as significant. The low number of participants in each group means that we cannot guarantee that these differences would be so clear in a larger sample or that they were representative of the population as a whole. Having said that there were some interesting patterns which emerged and which seem consistent with the qualitative feedback that we received from participants.

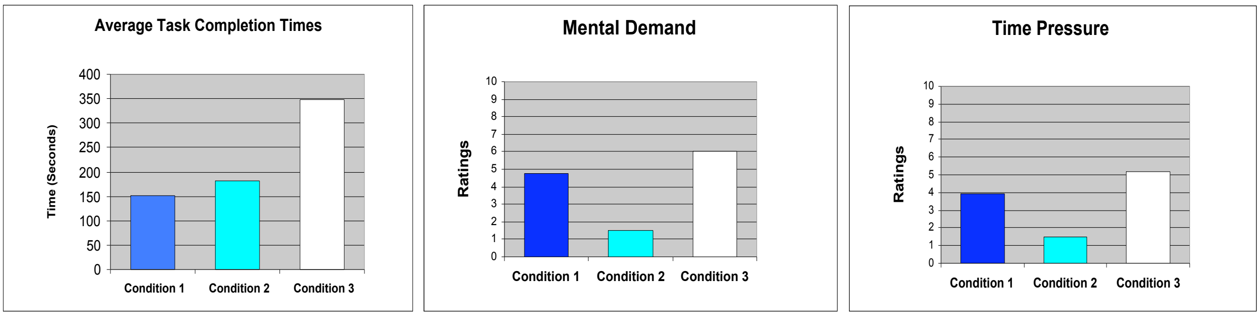


FIGURE 3: Graph showing average task completion times, mental demand and time pressure of participants.

Firstly, participants in conditions 1 and 2 completed the crisis management task (located the casualties) more quickly than those in condition 3. The qualitative data suggested that those in the third condition struggled when they came upon a point of confusion. Those in the first and second conditions did not, as they were able to question the HQ and were able to get answers to any questions that they had.

Participants seemed to feel that conditions 1 and 3 were the most demanding, were the hardest work and the most time pressured (Figure 3). This was perhaps not surprising given the different information provision and equipment involved in each. Although the levels reported were not that high they were clearly above those reported in condition 2. This seems to suggest that for some reason participants in this condition were more relaxed, and this could have been because of the nature of the interaction. In condition 2 participants did not have to study a map or use their visual system to interrogate the HMD screen; instead they were directed to their goal by voice command only. It is interesting that there was little difference in the workload levels between people asked to use the map presented on the HMD screen and those asked to use a paper copy of the map.

The situational awareness data are mixed and it is harder to see clear patterns. However participants in the second condition found it harder to get a handle on what they expected to happen in future, perhaps not surprising given they were largely reliant on good voice instructions to guide them around the building. Participants in other conditions could be said to have more information which they could interrogate themselves, and they had greater control over the presentation of information and the speed of the interaction in conditions 1 and 3. The findings show that those in condition 2 found it easiest to keep up with the flow of information.

Interestingly those in condition 1 reported having the best grasp of the situation and knowing what the best course of action to take. This suggests that despite the slightly higher reports of workload and demand, the extra information might have some value to participants.

Although it is not clear in the data, the only condition in which participants failed to complete the task was when they were not able to use the HMD. In either of the other two conditions errors in direction or map reading (getting lost) could be corrected by the HQ operator and by communication between the two. It would seem that the benefits of the HMD are not necessarily always obvious, especially when participants do not make errors. However, when errors occur or where the real world does not match the documented world (the map is out of date) the HQ and the HMD can be a useful support tool.

These data, whilst not conclusive suggest that further work might prove valuable in understanding the differences between workload demands when wearing the HMD and the balance that is required when it comes to not overloading the operator with information. In addition, control of information, resolution of uncertainty and awareness in the face of changing situations are other areas that can be considered.

5. CONCLUSION AND FUTURE WORK

5.1. Human Factors and socio-technical findings

In this section Human Factors issues that emerged during user testing the HMD kit are considered in greater detail.

Wearability and Mobility

Unfortunately one of the main findings was that users found that the HMD often slips and that it can be difficult to stabilise the view through the eye-piece. As the HMD rests across the forehead of the user it is likely that under stressful, hot or extended wearing, the HMD may slip even more as the user perspires. The wires that link the HMD to the laptop often get caught in clothing or under the user's arm. This could be resolved by securing the wires or building them into the user's clothing. The current set-up could result in the wearer having to use one of their free hands to hold the HMD in place during use, which clearly far from ideal. As these issues suggest, it is not clear whether the HMD kit will be comfortable to wear for long periods of time. When assembled, the full HMD kit is quite heavy and the current solution of a bag to carry the laptop is not ideal in terms of load distribution.

Wireless Connection

In an emergency situation gaining access to an existing wireless network might prove overly time consuming and technically difficult. The worst case scenario would be that passwords, account details and local network specialists would be a prerequisite to accessing the network. During trials the wireless network connection occasionally dropped out. These drop outs were particularly troublesome as re-establishing the network connection was found to be very difficult to co-ordinate. The field operator had difficulty reconnecting with the network, even when prompted by the HQ operator. This was mainly due to the limited amount of control that was afforded by the voice input medium. When network dropouts occurred the equipment had to be unpacked and reset manually. The difficulty of reconnecting was increased by the communications breakdown which was experienced during a network drop out. In addition, in certain locations wireless networks may struggle streaming so much data due to the bandwidth limitations. This needs to be tested more thoroughly.

Speech recognition

The concept demonstrator has a separate *voice/audio headset,* which conflicts with the headband. Partly for this reason, but also for robustness in noisy environments, alternative audio speech input or output devices (throat microphone or bone-conduction microphone) are being considered for future versions. Also, future trials may add background noise as an experimental variable. The software needs to be trained to recognise participants' voices. Therefore the experiment should be repeated with variable amounts of training in order to analyse the correlation between the voice training and accuracy.

Overview maps, buildings and floor plans

Building plans were not in a suitable format for a see-through display. A plain light background obscures the transmitted scene, so the background of a see-through information display should be dark, with text or other data presented in positive contrast. Valuable lessons were therefore learned about changes to building plans that may be useful, not only when using an HMD, but more generally enabling the reader to know better where they are with respect to the map. Also, out of date building plans cause inaccurate directional instructions. Building and floor plans need to be readily accessible on demand and possibly stored centrally by a governing body.

5.2. Future work

The future considerations involve mounting the display on a helmet rather than using a simple headband. In addition, development of the speech recognition part of the system (both hardware and software) should receive further attention. Alternative audio input and output devices such as throat microphone, bone-conduction microphone and earphone devices can be considered to improve the quality of speech in noisy environments. Other objectives include improving the usability of the wearable information system assembled and demonstrated by making use of more data (e.g. GIS, demographic) and also to improve the communication links of the demonstrator.

ACKNOWLEDGEMENTS

This work was partially funded by the European Commission through the OASIS project. (http://www.oasis-fp6.org/).

REFERENCES

[1] Carr, D. (1999) Guidelines for Designing Information Visualization Applications. Proceedings of ECUE'99, Stockholm, Sweden.
[2] Herdman, C. M et al., (2001) Attentional benefits and costs associated with integrating a Direct Voice Input (DVI) system into a multi-crew helicopter environment. Proceedings of the 11th International Symposium on Aviation Psychology, 2001.
[3] Shneiderman, B. (1996) The eyes have it: A task by data type taxonomy. IEEE Symposium on Visual Languages. Boulder, CO.
[4] McGuinness, B. and Foy, L. (2000) A subjective measure of SA: the Crew Awareness Rating Scale (CARS). In D. B. Kaber and M. R. Endsley (Eds). Human performance, situation awareness and automation: User centered design for the new millennium. Atlanta, GA: SA Technologies. (Paper presented at the Human Performance, Situational Awareness and Automation Conference, Savannah, Georgia, 16-19 Oct 2000.)

Detection and Tracking of Eye Blink to Identify Driver Fatigue and Napping

Indrachapa Buwaneka Bandara and Chris Hudson
Buckinghamshire Chilterns University College, CAC, HighWycombe, BUCKS, HP11 2JZ
i.bandara@bcuc.ac.uk, chris.hudson@bcuc.ac.uk

This work demonstrates the implementation of a high speed eye blink tracking system capable of detection and tracking of eye blink duration. Using a simple driving simulator, the performance in a driving situation can then be correlated to eye blink to provide a metric that can warn the driver when their sleepiness is becoming a danger to driving.

The system tracks eye blinks from the side of the eye. The reliable detection and tracking of eye blink is an important requirement for measure of PERCLOS (percentage eye closure) and blink duration in the detecting of driver alertness. Image acquisition and image processing algorithms are used for blink detection. By using a spectacle mounted sensor, the problem of analysis of head movement is minimised. We exploit the physiological properties and motion dynamics of eyelid movements (eye blink frequency, closure duration) to measure blink duration and PERCLOS. The three-step eye blink detection procedure: background estimation, template matching and tracking; techniques are used to analyse eye blink dynamics. To measure driver performance we designed a simple driver simulator using virtual reality tools in MatLab Simulink. The system is initially designed to allow the data to be stored for off-line analysis, but will eventually lead to a real-time detection system.

Motion dynamics, PERCLOS, blink duration, template matching

1. INTRODUCTION

Sleepiness increases the risk of a motorway crash either because the motorist falls asleep while driving or has reduced attention to driving. Many research and findings proved that falling asleep while driving is a major cause of road accidents in the world [1]. Detection of driver alertness and an ability to warn them before they reach a dangerous state of sleepiness is therefore important. In this paper, we mainly look at operational and technological methods of examining driver's eye blinking. John A Stem. [8] Pioneer of eye blink research says there is solid evidence that people's eye blink is related to tiredness or fatigue. Also he mentioned eye blink and eye movement tracking is the most reliable method for detecting human fatigue.

A few methods have been proposed for blink detection. In Yano et al. [3], frame differencing is used for blink detection. Al-Qayedi and Clark [4] track features about the eyes and infer blinks through detection of changes in the eye shape. Smith et al. [2] try to differentiate between occlusion of the eyes (due to rotation of the head) and blinking. The subject's sclera (the white of the eye) is detected using intensity information to indicate whether the eyes are open or closed (i.e., a blink is occurring). Black et al. [5] detect blink using an optical flow algorithm but the system restricts motion of the subject and needs "near frontal" views in order to be effective. The reported 65% success rate in detecting blinks [5] seems to be too low for driver fatigue or sleepiness detection. Generally eyes are tracked and correlation scores between the actual eye and the corresponding "closed-eye" template are used to detect blinks. Use of these methods is difficult to measure accurate blink durations. Another disadvantage of the system is that changing camera positions requires the whole system to be retrained. Furthermore, the same system may not be as effective if it were used on people of different races with disparate eye sizes and distance between the eyes.

The algorithm proposed here reduces complexity of the system by performing frame auto thresholding and template matching. However, problems still occur in detecting regions, due to various causes such as skin colour, eyeball motion, and possibly background illumination conditions. To differentiate between blinks and other types of motions, we use background estimation to detect the minor movements around the eye and generate a stabilized background for detection. Our goal is to create a system which (i) allows for different eye sclera (white segment of eye) section and (ii) intensity image analysis to reduce the lighting effect in detection. The advantage of using a head mounted camera is the removal of the need for head movement analysis.

The virtual reality tools in Matlab Simulink allow a real world environment for driving simulator to be produced to test driver performance by recording steering wheel movements, speed variations and reaction time. These parameters indicate driver alertness and compare the PERCLOS and eye blink duration to predict driver fatigue or sleepiness.

2. PROPOSED ALGORITHM

The blink detection algorithm uses background estimation, template matching and tracking of the eyes, by using image processing and image acquisition techniques.

2.1 Blink Detection

The steps in the blink detection algorithm are described below:

2.1.1 Background Estimation

(i) Estimate a stabilized background for eye tracking using the reshaping of images by converting a matrix input signal (144x176) to a row matrix, i.e., a 1-by-N matrix where N is 176.

(ii) The buffer input redistributes the input samples to a new frame size and generates a slower frame rate. The input frame period is Mi*Tsi, where Tsi is the sample period and Mi is input frame size. The output frame period is (Mo-L)*Tsi, which is equal to the sequence sample period when the Buffer overlap is Mo-1 where Mo is output frame size [6]. The output sample period is therefore related to the input sample period by:

$$T_{so} = \frac{\left(M_o - L\right)T_{si}}{M_o}$$

(iii) The majority of movements around the eye are due to vertical movements. The median value of each column (144 by 1) of the original matrix is calculated at each sample time to provide a M-by-1 vector (where M is 144) containing the median value for each column. This process helps to reduce the minor vertical movements to provide a stabilized image.

(iv) A threshold value is preset and any column with a median value below the threshold is set to all black. Columns where the median value is above the threshold are not changed. This background estimation process gives a stabilized output image for tracking.

2.1.2 Template Matching

(i) A Median Filter is then applied to the modified image using a 9x9 filter from a pre-established template to eliminate small areas of white. This provides a 2-D correlation.

(ii) A Morphological Operation Erosion process is then applied to remove the outer layer of the pixels from object. This process removes eye lashes.

(iii) The 2-D Correlation process computes the two-dimensional cross-correlation of two input matrices. Assuming that matrix A has dimensions (Ma, Na) and matrix B has dimensions (Mb, Nb) [6]. When the block calculates the full output size, the equation for the two-dimensional discrete cross-correlation is:

$$C_{(i,j)} = \sum_{m=0}^{(Ma-1)} \sum_{n=0}^{(Na-1)} A\left(m.n\right)conj\left(B\left(m+i, n+j\right)\right)$$

Where $0 \le i \le M_a + M_b - 1$ and $0 \le i \le N_a + N_b - 1$

Therefore the average size of the eye sclera template will correlate with the input image to filter outer disturbances around the original eye image.

(iv) A 2-D Maximum process then identifies the value and position of the largest element in input image and draws a rectangular box around that element.

2.2 Tracking of the Eyes

(i) Subtracting the background estimate output from the template matching output will give stable image with large elements. The absolute value of this resultant image will pass through an Auto threshold process to generate clear white binary elements.

(ii) Morphological closing performs a dilation operation followed by an erosion operation for the input binary image by using a predefined neighbourhood element to create segmented output for region filtering.

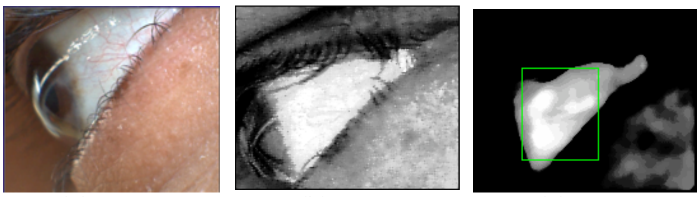
(iii) A Blob Analysis method calculates statistics for labelled regions (i.e. those surrounded by a rectangular box) in the binary image and returns the area of each element. The values are compared to a constant value to determine whether the input is within the specified constant range. The constant range is chosen to be from 2000 to 12000 pixels (this value was determined empirically).

This system captures 30 frames per second of 144x176 images. The average time taken for a complete human blink is about 300 to 400 milliseconds [7]. Each subject spends 20 minutes on the driving simulator. During this time 36,000 frames are captured. Using a 3.2GHz Pentium IV processor a real-time processing rate of about 22 fps can be achieved.

FIGURE 1: Complete system model of eye blink tracking system

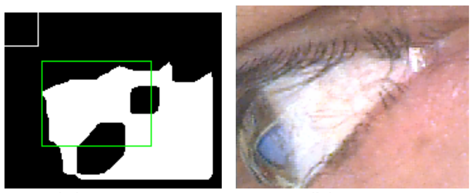
3. RESULTS AND DISCUSSION

Figure 1 shows the complete system of eye blink tracking. Figures 2 and 3, illustrate captured eye sclera sections in normal light and bright light conditions. The system clearly identifies the eye sclera as a binary image. Different light conditions change the eye sclera area; however the system is able to capture eye sclera in a specified pixel range. Using the background estimator reduces minor vertical movements around the eye. Median filtering with morphological erosion removes the eye lashes for clear eye tracking. Average size of eye sclera template minimizes the outer disturbances for eye sclera tracking. The captured area is within the rectangle and varies between the specified ranges of pixel sizes. A range of 2000 to 12000 pixels covers most of eye sizes.



(a) (b) (c)

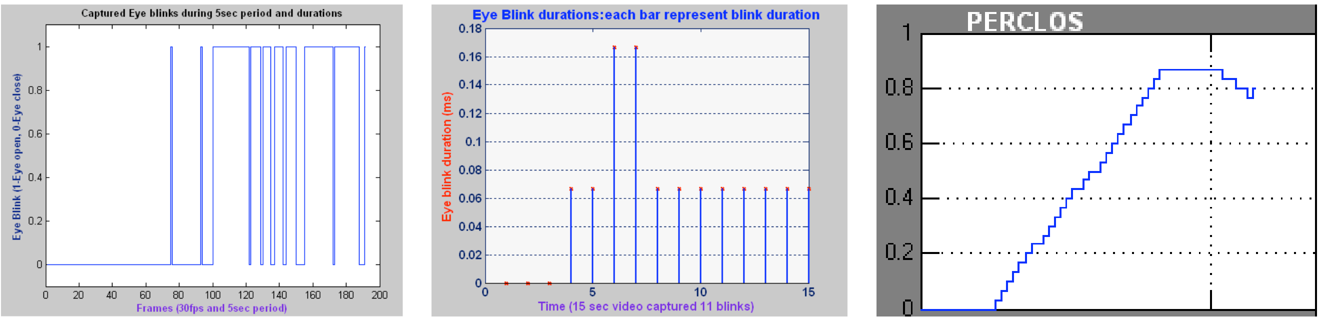
FIGURE 2: Result of eye tracking (a) Original Image
(b) Intensity Image before filtering
(c) Tracked eye sclera image,



(a) (b)

FIGURE 3: Different lighting condition
(a) captured eye sclera sector
(b) Bright light condition

The eye tracker gives three different outputs for data analysis – the number of blinks, the PERCLOS and the blink duration. Occasionally the system identifies more than one area within the 2000 to 12000 pixel range. The smaller area rectangle is then removed. The number of blinks and the total blink duration within a 1 second period is determined and passed to a moving average filter to measure PERCLOS. A high sampling rate gives accurate blink behaviour and fast response for rapid eye movements. Figure 4 (a) shows the eye blink capture where 1 is eye open and 0 is eye closed. In fact this gives a value of percentage eye open rather than percentage eye closed. Rapid eye blinks in short periods and slow eye blinks, lower the PERCLOS average and indicates the fatigue or sleepy conditions. This condition is then compared with the driver performance using the simulator. The intention is to provide statistics to correlate the PERCLOS or duration of blink with driver performance. The driver simulator records steering wheel variation, reaction time, average speed and checks every 10 seconds if driver goes out of the road limits. The algorithm was tested for video sequences of 12 different people, each of about average 20 minute.



(a) (b) (c)

FIGURE 4: Output data (a) captured eye blinks and durations (b) eye blink duration for each blink (c) Percentage eye closure (PERCLOS). Normal blink speed is 66ms shows in (b) and PERCLOS graph represent average value in between 1 and 0.8 is very alert state.

5-10 blinks were missed in 36,000 video frames and there were no false detections, giving an accuracy rate of 98% (there were an average of 25 blinks per minute. For 20 minutes testing this leads to an average of 500 blinks. If 10 blinks are missed during this period, this represents only 2%).

4. CONCLUSIONS

In this paper we have proposed an accurate and fast method for detection and tracking of eye blinks for predicting driver fatigue or sleepiness. Capturing of eye blinks using this method, the eye sclera section is extracted from the side of the eye. The head mounted camera reduces the complexity of the system by removing the need for head movement analysis. Once the blinking is detected, values such as blink duration, number of eye closures and PERCLOS can be calculated. The system was tested on 12 different subjects with an accuracy rate of 98% and more tests need to be made for different time periods of the day and different age groups. An initial questionnaire allowed such information as to the alertness of the subject at the start of the test, how much sleep they had the night before etc, to be determined. In addition a simple reaction test can be performed before and after the test. A standard test is used to determine the overall propensity to sleepiness of the test subject. Driver sleepiness or fatigue analysis is the main application of this system but it also can use for gaze-based human-computer interfaces and astronauts and pilots.

Future work involves the development of a real-time eye blink detector deployed on wireless (Bluetooth or WiFi) headsets used for mobile phone communication. The headset is easy to wear and measures driver alertness and warns them when danger occurs.

REFERENCES

[1] Journal article: Horne, J A and Reyner, L A. (1995) Sleep Related Vehicle Accidents. British Medical Journal, Vol.310, 1995, pp. 565-567.
[2] Journal article: Smith, P. Shah, M. and daVitoria, N. (2000) Monitoring head-eye motion for driver alertness with one camera, Pattern Recognition.
Conference:
[3] Yano, K. Ishihara, K. Makikawa, M. and Kusuoka, H. (1999) Detection of eye blinking from video camera with dynamic ROI fixation. *IEEE Trans. SMC*, Vol 6, pp 335-339,
[4] Al-Qayedi, A. M. and Clark, A. F. (2000) Constant-rate eye tracking and animation for model-based-coded video. *Proc. ICASSP'2000,*
[5] Black, M.J. Fleet, D.J. Yacoob, Y. (1998) A framework for modelling appearance change in image sequences. Proc. Intl. Conf. Computer Vision.
[6] MatLab Help, (2005) Version 7.1.0.246 (R14), August 02.
[7] Chapter: McWilliams. (1998) How fast does an eye blink. Medicine, University of Missouri- St. Louis, ID: 911393437.Me.
[8] Article: Stem, A. John. Fitzpatrick, Tony. (2004) The eyes have it: Candidates' eyes could be revealing, Centre for Security Technologies research studies alertness variables, (http://news-info.wustl.edu/news/page/normal/4018.html).

Blue Eye – Making Mood Boards in Augmented Reality

Jean-Bernard Martens, Andrés Lucero, Bart Naaijkens, Bastiaan Ekeler, Guus Rammeloo, Marcel van Heist, Matthijs Kwak, Max Sakovich
Department of Industrial Design, Eindhoven University of Technology,
P.O.Box 513, 5600 MB, Eindhoven, The Netherlands
{J.B.O.S.Martens; A.A.Lucero} @tue.nl

We describe the conceptual design of an augmented reality system for making mood boards. The main features of the proposed system are: the ability to integrate pictures of real objects and digitally retrieved pictures in an intuitive way, the ability to enrich mood boards by adding movements (and sounds), and the ability to control the system by means of hand gestures. The system can be realized by integrating available technology.

Interaction techniques, augmented tabletop systems, mood boards, design support

1. INTRODUCTION

Mood board making is a widely accepted and used technique in design processes [1]. Mood boards aim to capture the atmosphere of experiences by means of photographs or other expressive aids [2]. They provide a mechanism for designers to respond to perceptions about the brief, the problem as it emerges and the ideas as they develop [3]. Traditional mood boards are assembled by gluing different types of analogue media (pictures from magazines, photographs, fabrics, objects, etc.) on a mounting board. Digital mood boards can be created by collecting similar media in a digital format and assembling them with the help of graphics software, such as Photoshop, Illustrator or Freehand.

Within a student project at the Industrial Design department of our university we explored how mood boards could be created in augmented reality [4]. The design goal was to create a system that would combine advantages of the physical and the virtual (or digital) world. Six projects teams received the same project brief and worked simultaneously on alternative product designs. Within this paper we present the most promising product concept, which was created in the form of a video prototype. This prototype clarifies the major product characteristics and the relevant human-computer interactions, while paying less attention to the technologies required for realizing these functionalities.

After presenting the design process and the designed product, we discuss its relationship to existing work in human-computer interaction, in order to clarify both similarities and differences with existing designs. This relationship to existing work is also made to clarify that the product concept can be realized, based on available technology that has already been demonstrated elsewhere. The next obvious step is to actually build and deploy the system, in order to perform experiments with it that can inform a next design cycle.

2. DESIGN PROCESS

The six first-year industrial design (ID) students (authors 3-8) that were the members of the project team that created the conceptual design of the “Blue Eye” system did not possess advanced knowledge on how to create and use mood boards. In order to build up relevant expertise, they were first provided with a theoretical introduction on the subject by the second author, based on existing literature [2,3]. Under the supervision of experienced industrial designers who use mood boards in their design practice, the students were asked to individually create three types of mood boards: traditional, digital and in augmented reality. During the first week of the project, students created traditional mood boards by cutting out pictures from magazines and gluing them onto mounting boards. During the second week, they created digital mood board by retrieving images from the Internet and by using commercial software packages such as “Photoshop” for editing the pictures into a one-page composition. Last but not least, in the third week students created a mood board in augmented reality by using the Electronic Paper prototype (EPP) [5] which was implemented on an existing augmented reality system called the Visual Interaction Platform [6]. This EPP was designed to simulate, within a digital environment, early design activities such as sketching with pen on paper and arranging images. Its aim is to combine the naturalness of physical media with the flexibility of digital media. Although the EPP is not a mature design, in the sense that it is still undergoing improvements in terms of usability, it provided a means for students to gain hands-on experience with augmented reality.

For each of the three mood board creation techniques, students formulated what they perceived as main advantages and disadvantages.

Using “traditional” tools (i.e. scissors, glue) for making mood boards felt very natural, and the result remains (physically) available at all times. By its non-obtrusive presence, the mood board can remind designers of their own design goals, and invite others, such as clients or colleagues, to discuss it. The major obstacles were finding good input materials, such as suitable magazines, and keeping order of the mess of discarded material that easily arose in the workspace.

Using “digital” technology for creating mood boards provided access to a very large database of pictures (the Internet) and a wealth of editing functionalities. The main observed disadvantage was the difficulty of maintaining overview, both of available pictures and the global composition. Access to a suitable output medium, such as a print of sufficient size and quality, is required to make the mood board available to the designers and their environment in the remainder of the design process.

The EPP offered a large workspace which provided overview and in which (digital) pictures could be manipulated in a natural way. The fact that images were projected from the top onto the workspace was experienced as a drawback, since the hands were sometimes casting unwanted shadows. The functionality offered by the EPP was in no way comparable to that available in commercial software packages such as Photoshop, and some features, such as the ability to create layers, were duly missed. Some existing usability problems within the EPP, such as the observed latency between input actions and visual feedback, made it difficult to appreciate the full potential of the system. The use of the Internet as a resource for collecting images, and the need to print the resulting mood board for future use, were characteristics that were shared with the “digital” mood board.

2.1 Design Requirements

After analysis of the gathered experiences, it was decided to base the concept of a new mood board creation system on the following requirements:

- It should be possible to introduce physical pictures of arbitrary shapes, and pictures of actual objects, next to Internet pictures and digital photographs, into the digitally-stored and displayed mood board.
- The interaction should feel natural, in the sense that there should be a one-to-one correspondence between where the actions are performed and where the visual feedback is provided. This is inspired by the intuitive look and feel of the current augmented reality system.
- Instead of using external interaction elements, (two-handed) gestures should be used to control all operations. This would create a digital tool with an affordance that comes close to that of physical tools.
- The mood board prototype should provide functionality that can motivate the migration from analogue to digital media. Within the proposed prototype this was translated into the possibility of creating motion within the mood boards in a natural way. Other extensions such as adding sound were considered but are currently not included, in order not to overload the product concept.

3. SYSTEM DESCRIPTION

The hardware components of the system, see Figure 1, are a table with a large display surface, and two cameras that are mounted above the table, i.e., a low-resolution video camera for tracking hand gestures and a high-resolution digital camera for capturing still images (with newer digital cameras, see http://www.canon.co.uk/, both camera functionalities might be integrated into one camera). We currently use a video projector and a mirror underneath the table to create an image on a transparent plate, but other means for creating a large display surface might also serve the purpose. The table has a height of 90 cm (rather than 75 cm for a standard table) in order to allow for easy operation while standing up. Several people can easily gather around the table and cooperate in the mood board creating activity and associated discussion.

FIGURE 1: The Blue Eye system with the camera(s) mounted on top and the display integrated into the table surface.

There are different ways for creating images on the table. One is through digital means, i.e., by sending digital images to a dedicated “input” folder. Such images may for instance be acquired by a digital camera that sends pictures wirelessly to this folder (such as the Nikon Coolpix P1, see http://www.nikonusa.com/). The other way, see Figure 2, is by putting physical pictures or objects on the table and pressing the camera capture button. The back-projection allows using the well-known and developed Blue Screen technique (see *The Blue Screen Page* at http://www.seanet.com/Users/bradford/bluscrn.html) to capture images of the objects (and their outline) on the table. While a high-resolution image is captured, the projected background is changed into a uniform (and known) colour, so that the actual physical object can be easily segmented from the background (especially if we assume that we are looking for a single connected component on an otherwise uniform, or slowly varying, background). In case automatic segmentation is not to the user’s liking, or the user wants to extract a part of a digitally imported image, hand gestures may be used to cut out a (non-rectangular) region of an image.

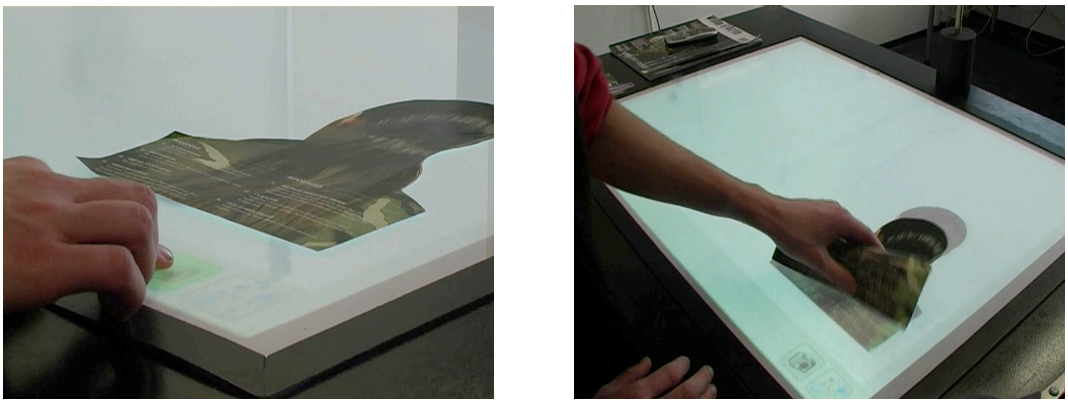


FIGURE 2: Images are captured in-place and a digital footprint remains after the physical object is taken away.

The Blue Eye system has a second mode of image capture that is set by a toggled button. The images that are captured while this button is active are considered to form a sequence and are interpolated to create a smooth and circular motion sequence (i.e., the last image of the sequence is linked to the first one). Most frequently, the sequence is created by capturing the same physical object in a number of positions.

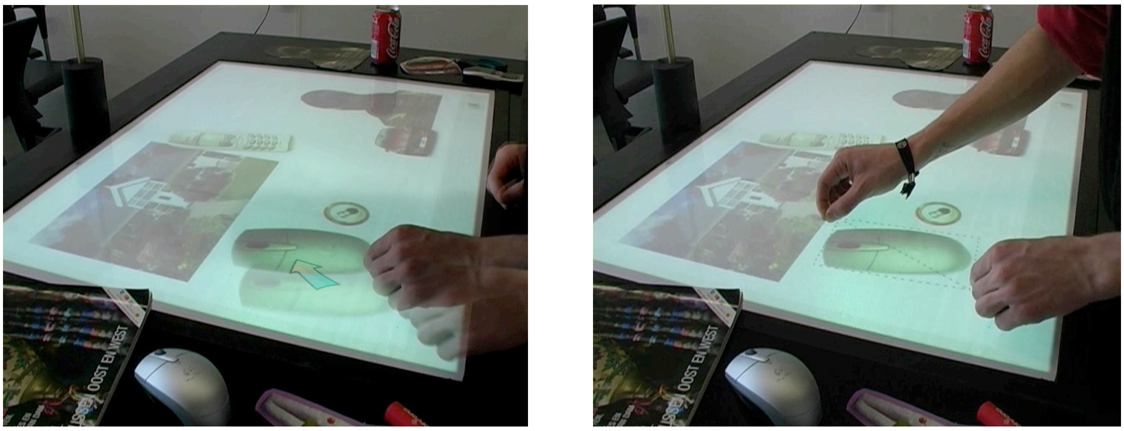


FIGURE 3: Image translating with one hand (left), and image resizing and rotating with two hands (right).

A video camera (i.e., a firewire camera with a spatial resolution of 640x480 pixels, capturing 60 frames per second) is used to track the gestures that are needed for operating the system. An obvious advantage of optical tracking is, amongst others, that several users can interact at the same time, using both hands. The interaction gestures could either be executed by means of tagged interaction elements, which allows for a robust and easy solution from a technology perspective [5], or using human hands, which is obviously the most flexible solution from the point of view of the end user. The feasibility of performing real-time tracking of arm and hand motions, and of using them for human-computer interaction, has been demonstrated earlier [7], and some software solutions are publicly available (see http://ilab.cs.ucsb.edu/projects/mathias/handvu_ilab.html). The requirements for the Blue Eye system are limited since only few gestures need to be distinguished. For instance, closing the index finger and thumb (of one hand in case of translation; of both hands in case of resizing and rotation) could signal a re-positioning event to the system, as shown in Figure 3. Translation and rotation might also be integrated in a single-hand gesture using the technique proposed in [8]. An extended index finger might be used to draw on the display, while an extended middle finger might be used to cut out parts of an image. Such hand gestures are expected to be natural and easy to learn.

The mood board output is an image that can at any stage be retrieved from a designated “output” folder and imported into other applications.

4. RELATIONSHIP TO EXISTING SYSTEMS AND FUTURE WORK

The idea of capturing information from the real world in an intuitive and compelling way has also been expressed in other recent designs [1,9]. The “I/O Brush” [9] is a paintbrush with an integrated camera, light source and touch sensors that can pick up colors, textures and short motion sequences from its environment. The physical size of the I/O brush limits the physical elements that can be “picked” to object details rather than complete objects. These picked elements are used as brushes for drawing, hence creating interesting and compelling pictures. A similar functionality is feasible within the proposed system. An important difference is that both larger objects and object details (by cutting out a region of a captured object) can easily be captured in the Blue Eye system. The possibility of creating motion sequences from a series of captured still images is another obvious extension.

The “Cabinet” prototype is an augmented reality system for managing photo collections [1]. It can photograph physical objects on the table surface and replace them by a digital footprint in place. The captured picture however represents the entire workspace and only a rectangular region of interest can be specified by the user to crop the image. Interactions within the Cabinet system are performed by means of a pen on a digital tablet, which has the consequence that images need to be projected from the front, rather than from the back, as is done in our case.

Our own Electronic Paper prototype (EPP) [5] demonstrates how image handling and drawing can be used so support early design activities. The EPP has much more extensive functionality than the proposed system, but uses only digitally generated images. It does not possess the natural and seamless interaction with physical objects that is proposed in the Blue Eye system. The use of a digital pen and the physical construction of the EPP also limit it to be single user.

The next step is to actually implement and test the system. The picture capturing requires a computer-controlled digital camera (such as Canon PowerShot S80 with http://www.breezesys.com/PSRemote software). Motion tweening (morphing), i.e., creating image frames in between specified key-frames, is available, for instance in commercial packages such as Flash. The biggest obstacle to the system realization is most likely the hand tracking and gesture recognition. As an intermediate solution, the available tracking technology within our VIP system [6], based on using an infrared-sensitive camera and infrared-tagged objects, can be used. This might be implemented by mounting infrared-reflecting dots on both hands of the user. Using this technology, we can start gathering end-user experiences before we have gesture recognition available.

Once the system has been realized, the extension with additional functionality can be explored. One potentially interesting idea arose in one of the other student projects. They proposed to assign colours to different parts of a soundtrack and to associate (for instance by means of dragging) the different colours to different parts of the mood board. Not only can this be used to add sound to a mood board, but it also provides a means for sequencing through the different components of a mood board. Of course, such interactions need to be tested in order to decide whether or not they are truly useful.

REFERENCES

[1] Keller, I. (2005) For Inspiration Only: Designer Interaction with Informal Collections of Visual Material, Ph.D. thesis, Technische Universiteit Delft.
[2] Muller, W. (2001) Order and Meaning in Design, Lemma Publisher, Utrecht, the Netherlands.
[3] Garner, S. and McDonagh-Philp, D. (2001) Problem Interpretation and Resolution via Visual Stimuli: The Use of Mood Boards, The Journal of Art and Design Education, 57-64.
[4] Milgram, P. and Kishino, F. (1994) A Taxonomy of Mixed Reality Visual Displays, IEICE Transactions on Information Systems, 1321-1329.
[5] Aliakseyeu, D., Martens, J.B. and Rauterberg, M. (2006) A Computer Support Tool for the Early Stages of Architectural Design, accepted for publication in Interacting with Computers, see also: http://authors.elsevier.com/sd/article/S0953543805001128.
[6] Aliakseyeu, D., Martens, J.B., Subramanian, S., Vroubel, M. and Wesselink, W. (2001) The Visual Interaction Platform, in: Proceedings of Interact 2001, IOS Press, 232-239.
[7] May, R. (2004) Towards Directly Mediated Interaction in Computer Supported Environments, Ph.D. thesis, University of Washington.
[8] Kruger, R., Carpendale, S., Scott, S.D. and Tang, A. (2005) Fluid Integration of Rotation and Translation, in: Proceedings of CHI 2005, ACM Press, 601-610.
[9] Ryokai, K., Marti, S. and Ishii, H. (2004) I/O Brush: Drawing with Everyday Objects as Ink, in: Proceedings of CHI 2004, ACM Press, 303-310.

Sketch Tool Usability: Allowing the User to Disengage

Beryl Plimmer, Gene Tang and Mark Young
University of Auckland, Private bag 92019, Auckland, New Zealand
beryl@cs.auckland.ac.nz, mark@educk.co.nz, gene_tang@hotmail.com

The key to a successful early design environment is an unobtrusive user experience. Yet there has been little focus on the usability of sketch tools. Here we describe two iterations of prototype and usability evaluation of InkKit, a domain independent sketch-tool for diagramming. We found that reducing constraints, careful use of metaphor and pairing input and output devices improved the user experience.

Sketch tools usability, pen-based interaction, metaphor

1. INTRODUCTION

Diagrams are used in many disciplines to visually describe ideas. When first creating a diagram, the ability to quickly express ideas ambiguously and incompletely is important [4]. As this is not possible with most computer-based tools, designers do not use them until their design is almost complete. Instead they favour traditional tools, paper, whiteboards, pens and erasers. Studies of computer widget-based design environments suggest that the functions that interfere with the design process are: the requirement to select a specific widget; distractions cause by tidying the artefact while designing (aligning and sizing); and the formal appearance of the diagram – which implies completeness [4, 10]. To address these issues computer-based sketch tools closely emulate traditional tools in that they afford pen input of informal diagrams but add value by providing computational support for editing and archiving of artefacts. Additional benefits can ensue with intelligent sketch recognition for example automatic beautification, execution and translation of the hand-drawn sketches. Yet, if this functionality becomes obtrusive the benefits of an informal environment are lost. The challenge is to provide additional functionality within the sketch-tool's user interface while minimising its cogitative demands so that the designer can disengage from operating the computer and therefore fully engage in the design process.

2. BACKGROUND

The central function of a sketch tool is to support digital inking on to a digital canvas. Normal computer editing is also standard: we experimented with not providing editing (except erase) with Freeform [11], however our first usability test showed users expect and value computer support for editing. Most sketch tools also include recognition engines to enable execution and translation of the sketch. However many restrict recognition to drawing shapes – thus requiring keyboard input of text [6, 7], this is a distraction for the user [5]. Those that support hand-writing recognition are very constrained [8] or require the user to change modes [12].

Two approaches to displaying sketches are evident in the literature; an infinitely large canvas or many small canvases. A large canvas presents an interaction challenge as only a portion of the drawing can be viewed at normal resolution at any one time. To support navigation these tools have either a radar window [3] or zooming [9], another, unexplored possibility is a fish-eye view. Small canvases are restricted in size so that they can be viewed and edited at full resolution. They are often accompanied by a storyboard [1, 7, 12], where the collection of canvases can be shown as thumbnails. The storyboard is a linear sequence of sketches where the user can establish relationships between the sketches by either placing them in a particular order or establishing links between the sketches.

The functionality that makes computer-based sketch tools useful also complicates the user interaction. In particular mode and input device changes may be necessary to move between drawing and writing, inking and editing and storyboard and sketch. These mode changes are different to those in a tactile environment where one may change tools or physically move objects. Problems arise when it is not clear to the user what mode they are in and what actions are available. The drawing/writing mode changes in FreeForm [12] cause these types of interaction problems. Like other drawing environments, most sketch tools require a mode change for editing. Once in this mode ink is selected and then changed either via icons and grab-handles or functional gestures. Menus and icons provide certainty of operation, whereas functional gestures are less reliable and, if recognition fails, create more work for the user.

Pen or stylus input devices are used for sketch tools. Digital whiteboards are useful when a large space is desirable for collaborative work. However the input data that digital whiteboards provide is of a lower quality than that of digitizer tablets, this makes accurate recognition of ink more difficult. Earlier systems used input digitizers such as Wacom™ tablets. These provide accurate position, time and pressure data but are disconnected from the output visualization which is usually on an adjacent display. Tablet PCs and newer Wacom tablets provide similar input accuracy with the advantage of the input and output surfaces being one and the same. In the following section we describe our experiences designing, building and evaluating the user interface for InkKit, a generic diagramming environment.

3. OUR APPROACH

InkKit is a domain independent, extensible diagramming environment. The goal is to provide a context free environment that imposes no additional cognitive load on the user: basic designing must be as simple as pen and paper. Yet InkKit is extensible so that it can incorporate sophisticated support for beautification, execution and translation of diagrams. InkKit runs on the Microsoft Tablet OS utilizing the tablet hardware and OS character recognition. As with any computer-based drawing tool, display space is a constraint. This is exasperated with the tablet OS as digital whiteboards are not yet supported as input devices. Apart from the user interface the other major part of InkKit is the recognition engine. InkKit is the first sketch tool to provide comprehensive, modeless drawing and writing recognition. Adaptation for specific types of diagrams is achieved by providing example components to the recognition engine. Extensibility allows user written modules to export recognized diagrams into different formats. We have implemented a variety of domains including user interface diagrams, organization charts and UML class diagrams. Output formats include html, java, Microsoft Word drawing objects and a variety of picture formats. A report has been published on the development of the first prototype [2]. Here we discuss the user interface of that prototype, the usability evaluation of it, and the subsequent reengineering and retesting

The design of the first prototype was based on our experience with FreeForm [12] and the experiences of other researchers. With the user interface we experimented with a single view comprising of multiple MDI forms (figure 1), removing the constraint of linearity and fixed sized storyboard elements evident in most sketch tools. There are two modes; storyboard and form, to change mode the user taps an icon on the tool bar. To visually differentiate modes, in storyboard mode the ink is greyed, in form mode the ink on the active form is black. In storyboard mode the forms are manipulated as entities, they can be positioned anywhere on the storyboard and resizing zooms the sketch. Relationships can be established between forms by creating a link by first tapping on the start-point of the link and then on the end-point. The meaning of the link is domain dependent, for example a user interface form with a link between a dropdown control and a form with words on it indicates that the list should be filled with the words. With organization charts the links may indicate different departments. In form mode the sketch is a computer supported drawing space. Unlike other sketch tools InkKit provides a modeless drawing/writing sketch space; this means, that users can write and draw on the form much as they would on a piece of paper. To edit ink users must first enter edit mode by tapping the lasso icon on the toolbar, this changes the cursor and stylus contact with the surface produces a dotted line. Selected ink is surrounded by a bounding box, dragging this boxes handles resizes the ink. Undo and redo are activated by tapping the appropriate icon. Resizing a form in form-mode changes the page size. We have not incorporated functional gestures into InkKit as our experiences with FreeForm suggested that they are not intuitive.

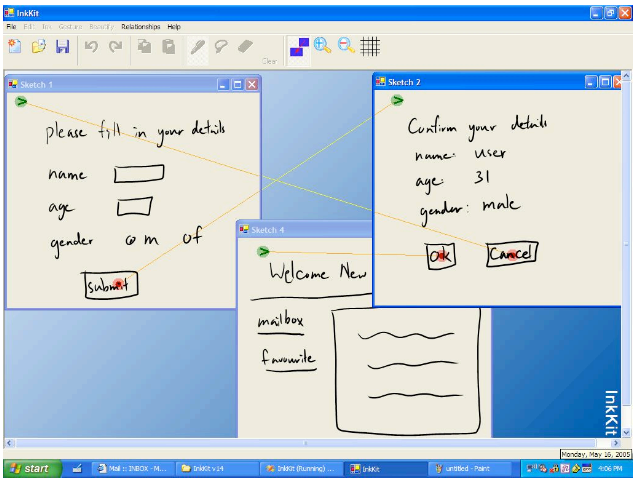


FIGURE 1: InkKit first prototype single-view user interface

FIGURE 2: InkKit second prototype two display, dual-view user interface

3.1 Usability Evaluation

InkKit is the first sketch tool specifically designed as a domain independent environment. This usability evaluation checked the fundamental design premises, such as the interface metaphor and map between the system model and user model, and also detailed interaction techniques.

The terminology used to describe an interface is usually metaphorical and is a powerful suggestion of the system model. This first prototype of InkKit adopted a mixture of computer and design jargon commonly used for sketch tools. As the goal is to provide a domain independent sketch tool we re-examined the terminology and compiled a short list of possible metaphors and associated words including: parent form, storyboard and portfolio and child form, sketch and canvas. We then showed InkKit to a variety of people and asked them which they thought best described it. We found that parent and child form were too computer specific, storyboard was not widely understood and that it implied a linear sequence of sketches (as used in film storyboarding), and canvas is more closely associated with the fabric than an artist's canvas. The best metaphor is a *portfolio* full of *sketches*; hence we changed the terminology to portfolio and sketch.

We also conducted a formal usability test of this prototype on at Toshiba M200 tablet. The six test participants were 4th year software engineering students. As the participants were unfamiliar with the tablet PC and InkKit a short training session was given first. Then each was asked to create a three-form user interface for joining a sports team supporter's website. The sessions were observed and recorded using Morae™ usability testing software. The sketch space and mode was well received by the users. The main difficulty we identified with our previous tool Freeform was the mode change required for drawing and writing; we have eliminated this in InkKit with a more sophisticated approach to recognition. Editing support for ink was commented on as an advantage of digital sketch tools. And although this requires a mode change, the change of cursor, and appearance of the stylus path were sufficiently clear to users for this not to cause difficulties.

However the storyboard mode was not as successful. Moving between form editing and storyboarding required a mode change. Just as we had found with inking modes with Freeform, this created real difficulties for the users. Although InkKit greyed the ink on the forms to indicate storyboard mode users frequently attempted form editing functions in storyboard mode and storyboard functions in editing mode. We also identified an interaction problem when creating links between forms (taps to position start and finish points); most users tried to drag a link. In addition to these problems, the small display size of the tablets means that a multi-form design either has a large percentage of the forms obscured or the forms zoomed very small.

3.2 Reengineering

To recap: we found that the terminology we had adopted from previous sketch tools was poorly understood. We also identified three major usability problems; the mode changes between storyboard and form, limited display space and the link creation method.

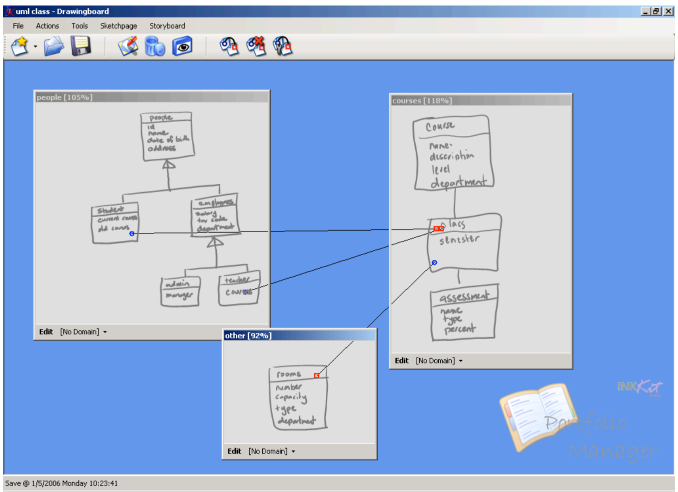


FIGURE 3: InkKit second prototype portfolio

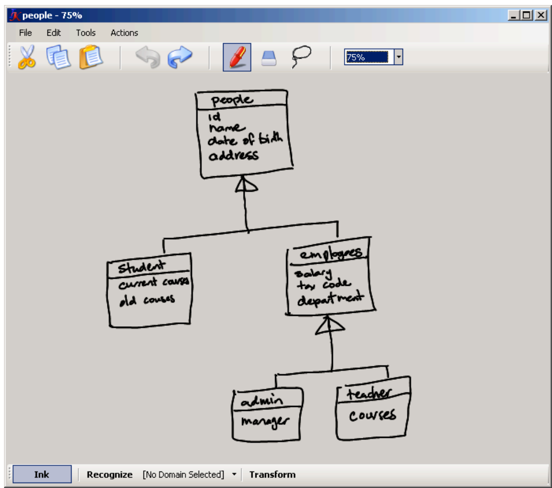


FIGURE 4: InkKit second prototype sketch

In principle the concept of a single view is appealing, however two of the usability problems, modality and display size are linked to this approach. We considered various approaches to reducing the impact of modality however none were likely to alleviate it completely. The display size limitations are solvable by networking tablet computers or using e-whiteboards. However networking computers adds a layer of complexity and is quite clumsy and e-whiteboards are not yet of sufficient quality to support accurate recognition. We chose to adopt a two monitor, two view interface; portfolio and sketch (Figure 2).

The portfolio (Figure 3), (the storyboard mode of the previous prototype), is displayed on an auxiliary monitor; each MDI form (sketch) has an edit button added to the bottom left of the form and the ink is always greyed. The link creation method was altered so that the user drags a line from the start-point to end-point. To edit a sketch the user opens the MDI form as a separate window (Figure 4). The sketches are shown on the tablet screen (where the stylus can be used for input). A number of architectural changes to the software were required so that the two views are synchronized. If no auxiliary display is available the portfolio is shown on the tablet display as an independent window.

We usability tested InkKit again using the same test setup described above with different test participants. Users commented that it was clear that inking was not available in the portfolio view and that it was easy to swap between windows. Creation of links by dragging also solved that problem. However the auxiliary display created a different problem: a natural split of input devices is to use the stylus for the tablet and mouse for the auxiliary display. However, because the OS links the mouse and tablet cursors, each time a user moves focus to the other display both cursors follow. This causes a particular problem with the auxiliary display as manually moving the mouse cursor back to the display is slow. We are modifying InkKit so that when it is running with two displays the cursors and pointing device input events are split – the stylus cursor and events are handled by the active window on the tablet and the mouse cursor and event are handled by the active window on the auxiliary display.

4. DISCUSSION AND CONCLUSION

Early design environments must allow quick unconstrained recording of ideas. To provide this in a computer-based environment the computer interaction must be context free and minimize cognitive load so the designer can fully engage in the design process. Yet the value of computer-based sketch tools is derived from their added functionality. The initial design of the InkKit user interface was based on our experience with Freeform and a survey of other sketch tools and hardware options. Our first prototype was implemented with a single view interface requiring mode changes to indicate whether actions were to the storyboard or an individual sketch. The evaluation of this prototype suggested that different terminology provided a better metaphor and user testing showed that the single view modal interface and limited display space was likely to cause ongoing interaction problems.

We re-engineered the interface, changing the terminology and splitting the interaction between two views, the portfolio, and individual sketches. The portfolio is displayed on an auxiliary monitor, thus maximizing viewing space. Retesting showed that the new interface solved the modality problems while increasing the viewable area. Our original approach of a single viewable space remains closer to a tactile environment such as a tabletop or whiteboard and may be worthy of re-examination as large displays with sufficient input accuracy emerge. InkKit, in its current form, provides and excellent, easy to use, context free sketching interface that is supported unobtrusively by a powerful recognition engine.

An easy-to-use sketch tool allows the designer to engage in the design process rather than the computer interaction. Ensuring that sketch tools are intuitive requires ongoing careful design and evaluation as new hardware and more intelligent support continues to change the user environment.

REFERENCES

[1] Bailey, B. P. and Konstan, J. A., (2003) *Are Informal Tools Better? Comparing DEMAIS, Pencil and Paper, and Authorware for Early Multimedia Design*. Proceedings *CHI 2003*. Ft Lauderdale: ACM.

[2] Chung, R., Mirica, P. and Plimmer, B., (2005) *InkKit: A Generic Design Tool for the Tablet PC*. Proceedings *CHINZ 05*. Auckland: ACM.

[3] Damm, C. H., Hansen, K. M. and Thomsen, M., (2000) *Tool support for cooperative object-oriented design: Gesture based modelling on and electronic whiteboard*. Proceedings *Chi 2000*: ACM.

[4] Goel, V., (1995) *Sketches of thought*. The MIT Press: Cambridge, Massachusetts.

[5] Jarrett, R. and Su, P., (2003) *Building Tablet PC applications*. Microsoft Press: Redmond.

[6] Kara, L. B. and Stahovich, T. F., (2004) *Sim-U-Sketch: a sketch-based interface for SimuLink*. Proceedings *of the working conference on advanced visual interfaces*. Gallipoli, Italy: ACM Press.

[7] Landay, J. A., (1996) *Interactive sketching for the early stages of user interface design*. Carnegie Mellon University: Pittsburgh, PA.

[8] LaViola Jr., J. J. and Zeleznik, R. C., (2004) *MathPad2: a system for the creation and exploration of mathematical sketches*, in *ACM Trans. Graph.* p. 432-440.

[9] Lin, J., Newman, M. W., Hong, J. I. and Landay, J. A., (2000) *Denim: Finding a tighter fit between tools and practice for web design*. Proceedings *Chi 2000*: ACM.

[10] Plimmer, B. and Apperley, M., (2005) *From Sketch to Blueprint: Supporting the creative design process*. Proceedings *Workshop on improving and assessing pen-based input techniques*. Edinburgh.
[11] Plimmer, B. E. and Apperley, M., (2002) *Computer-aided sketching to capture preliminary design*. Proceedings *Australasian User Interface Conference*. Melbourne.
[12] Plimmer, B. E. and Apperley, M., (2003) *Freeform: A Tool for Sketching Form Designs*. Proceedings *BHCI*. Bath.

Enhancing Web Accessibility Through an Adaptive System

ChuiChui (Samantha) Tan, Wai Yu, Graham McAllister
Virtual Engineering Centre & Sonic Arts Research Centre, Queen's University Belfast, Belfast, BT9 5HN, UK
{s.tan, w.yu, g.mcallister} @ qub.ac.uk

In the age of the Internet, we have instant access to enormous amounts of information. However, due to the inaccessibility in the technology itself, people with visual impairments are not permitted full advantage of this service. Hence, research has been carried out to improve Internet accessibility for visual impaired users. Nevertheless, most of the research has been developed to solve a specific problem based on a particular criterion. This project aims to develop an adaptive architecture which brings various existing assistive technologies together into a coherent system. This system is designed to be interoperable so that (1) it can support various assistive devices; (2) it is capable of dealing with diverse types and formats of graphical web content; and (3) it takes consideration of user's profile and preferences. This paper firstly gives an overview of the system, along with background information and related approaches. The main part of this paper will describe in detail the launching of the adaptive system that has been developed, including its design considerations and working principles.

visually impaired, graphical web content, web accessibility, adaptive architecture, interoperable, assistive technology

1. INTRODUCTION

With the advent of the Internet, we can now easily gain access to vast amounts of information and services online. While we are enjoying the benefits of the Internet, people with visual impairments have been excluded from full access to this technology by the associated inaccessibility issues. One of the major hindrances is caused by the use of graphics which are not accessible to a screen reader [1] or a Braille display. These devices represent two of the most common assistive technologies used by most of the visually impaired people when they are surfing the Internet [2]. Consequently, studies have been carried out to improve Internet accessibility of visually impaired people by substituting the visual modality with audio and haptic feedback. For instance, O'Malley et al. [3] have developed a system that simplifies the authoring of 3D haptic web content so that this content is more accessible to visually impaired people. Meanwhile, a system accomplished by Way and Barner [4] is able to convert bitmap images from visual to tactile form – the TACTile Image Creation System (TACTICS) which provides access to visual information. All of these studies have brought a great advantage to those with visual disabilities; however these research efforts focus on a particular problem and only work with a specific set of hardware. There is a lack of an integrated platform that allows visually impaired people to access any graphical web content without worrying about the type of device to use. This paper describes the ongoing development of an adaptive system that is believed to be able to solve the aforementioned problems. It aims to achieve this by incorporating different technologies to increase the accessibility of graphics for visually impaired users and at the same time, taking care of the user's preferences. The launching of the system has already been developed and its detailed description, design principles and working environment form the core of this paper.

2. ARCHITECTURE DESCRIPTION

This work is part of an ongoing European project, ENABLED (see Acknowledgements). One of the main objectives of this research is to develop an adaptive system which is able to create an adequate interface for people with sight difficulties. AMIS [5], a project developed by Kawamura et al which holds similar objectives, employs a flexible XML-based architecture that allows for adaptation of the standard interface to meet the needs of both users and assistive technologies. Meanwhile, AVANTI system developed by Stephanidis et al [6] uses adaptability and adaptivity techniques to provide high-quality interaction to users with different abilities, skills and preferences. The adaptive system described in this paper is developed to be interoperable by supporting multiple assistive devices, independent of graphical web content and adaptable to the user's profile. As there is an extensive range of assistive technologies available to the users, it is essential to take this issue into account as one of the main elements in the development of this system. By categorizing the types and formats of graphics available on the Internet, it can simplify the process of making graphical web content accessible to the visually impaired. For example, one of the graphic types that we can easily find on the web is graphs, and the file format for a graph can be in JPEG or PNG (for bitmapped graphs), SVG (scalable vector graphics) or even in Excel form (spreadsheet graph). Meanwhile, according to Holland et al. [7], mining user preferences has become one of the key focuses considered by a lot of researches. As a result, several techniques to build user-adaptive web sites have been developed in recent years. In this project, we regard that the description of users' capabilities and preferences is necessary in this adaptive system. A simple illustration of the system's concept is shown in Figure 1. This adaptive system can work with two of the most popular Web browsers – Microsoft Internet Explorer and Mozilla Firefox.

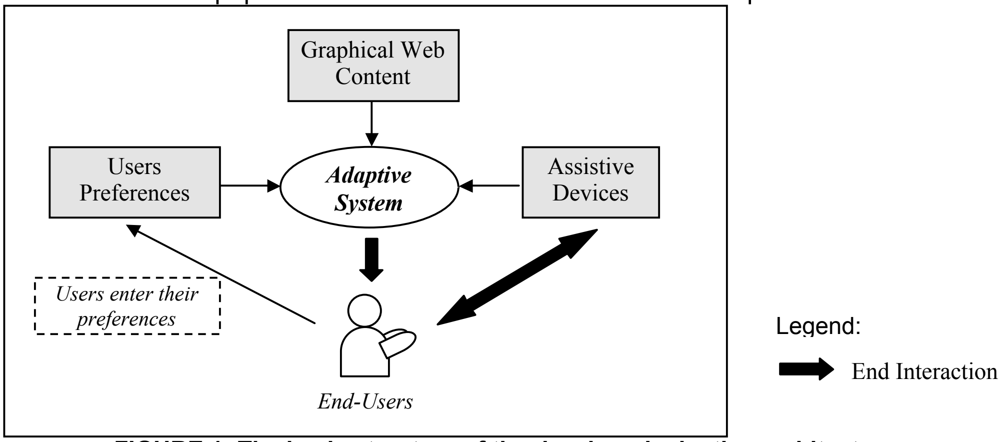


FIGURE 1: The basic structure of the developed adaptive architecture

In order to implement this system, as illustrated in Figure 2, sub-systems have been built to handle and integrate the three main variables (as shown in Figure 1). All of them play a vital role in the system and interact with each other. The Web Content sub-system detects the type of web content explored by the user, as well as its format. Through the Configuration sub-system and Preference sub-system, the adaptive system is able to automatically detect the user's preferences and the assistive technologies available on the user's machine. Additionally, the user is also allowed to change their preferences setting anytime and save it in the Preferences sub-system. Sub-applications are small independent applications that essentially translate the graphical content into an internal representation. This can then be used to create a geometric representation of the graphical content such as curves, surfaces and so on. This geometric representation can be used to get the corresponding haptic and audio feedback by analysing the user's cursor position. This paper describes the Web Content sub-system that has already been developed and shows how it integrates with a sub-application for graphs.

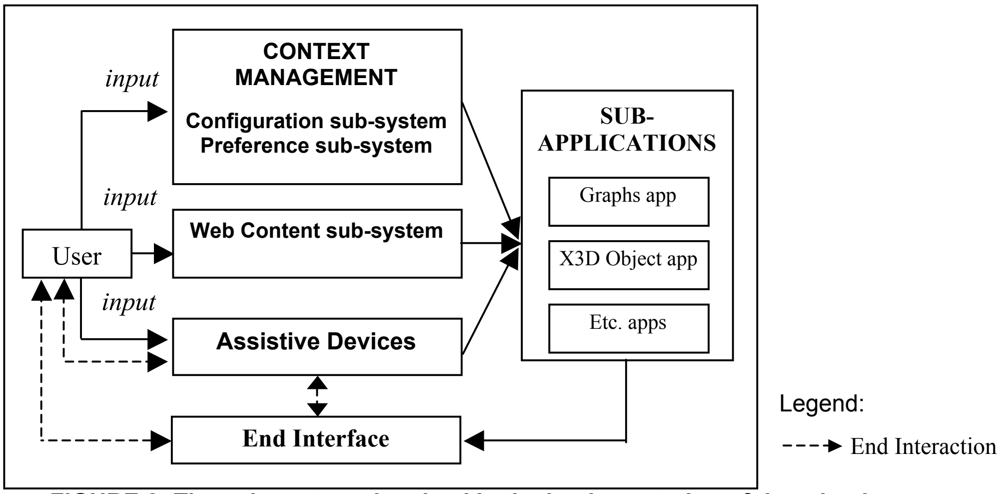


FIGURE 2: The sub-systems involved in the implementation of the adaptive system

3. WEB CONTENT SUB-SYSTEM

The Web Content sub-system is one of the main parts in the adaptive system. It is based on the Microsoft Active Accessibility (MSAA) technology, which is used on the Windows operating system to provide a standard way to get information about user interface elements. The Web Content sub-system holds two main roles: (1) to launch the sub-application according to the type of the graphic the user wishes to explore; and (2) to download the data file of that particular graphic to the local machine so that it can be accessed by the sub-applications. The data file is the file that contains raw data of a graphic which is required by the sub-application in order to create a suitable interface to the user. Collaboration from Web developers is needed so as to allow the user to use this system to explore the graphics on their Web pages. The Web developer only needs to add a button with a specific name near to the graphic and provide the URL of the data file as the title of the button. For example, if the graphic is a spreadsheet graph, then the data file should be an

Excel file which consists of the data and information about the graph. Each of the graphics will have its own button as this button carries the data file belongs to the graphic.

3.1 Design factors

In the process of the design and development of the adaptive system, it is important to consider factors that might affect the end result. There are several issues which are regarded to be significant in this system. The way the system interacts with the end-users should be as straightforward as possible, for instance, a complicated concept to launch the application ought to be avoided. As the target user group is visually impaired people, alternatives to meet the purposes of this Web Content sub-system are proposed according to the requirements and restrictions faced by the users. For example, it is a good practice to have the least pop-up boxes prompted to the users as some of the screen reader versions are not able to read the content of these message boxes. They will only read the title bar of the message box and this may cause problems for the users. Furthermore, it is essential to consider the common approaches users with visual disabilities use computers and the Internet. Although there are a number of haptic and tactile mice introduced at the present to help visually impaired users to access the Internet, there are still a lot of users, especially congenitally blind users, who are not so comfortable when using the mouse [8]. Hence, having functions that can be controlled by both mouse and keyboard is one of the main considerations while developing this system.

Another major design concern is the security issue of this Web-based application. As Internet users are more aware of the Web application security vulnerabilities, most of the Web browsers are configured to provide increased security functionality to protect their users. There are a few ways to launch a Web-based application. Some of them require the user to lower the security setting on their computer, whereas others could prompt a few security notifications on the screen to ask for the user's agreement to install and run the application even though the user has the default security setting in their browser. These scenarios are not encouraged. For the former condition, the users might not agree to lower the security setting of their browsers; whereas for the latter, it might cause some confusion and inconvenience to the visually impaired users by having too many message boxes prompted to them.

3.2 Start-up procedures

To be able to use this adaptive system, the user needs to install the application onto their computer in the first place. Additionally, the user has to be aware that, a Web page which is compatible with this adaptive system has one specific button containing the word "ENABLED" for each graphic. This button acts as the "start switch" of this system in which it listens to the user action when they are ready to explore a graphic using this adaptive system (ENABLED system). Most of the visually impaired users who use the Internet in daily life have a screen reader installed on their computer. As screen readers are able to read out the text displayed on buttons, the user can easily know the existence of the ENABLED button on the Web page. Since every button carries the data file of each graphic and there could be more than one graphic on a Web page, the user needs to ensure that the keyboard focus or mouse focus is on the button which is near to the graphic they are interested in. This can be accomplished through their screen readers. Subsequently, the user informs the system that they are ready to explore the graphic in more detail by pressing a key in the keyboard. In the current implementation, the designated key is the *ENTER* key. Through features provided by MSAA technology, the application is able to know the moment the user is on the specific button and is pressing the specific key. Consequently, the data file of that graphic will be downloaded and the related sub-application will be launched automatically by referring to the type and format of the graphic. From this point onwards, the user can then start to explore that particular graphic with the assistive devices they own. The steps are portrayed in Figure 3.

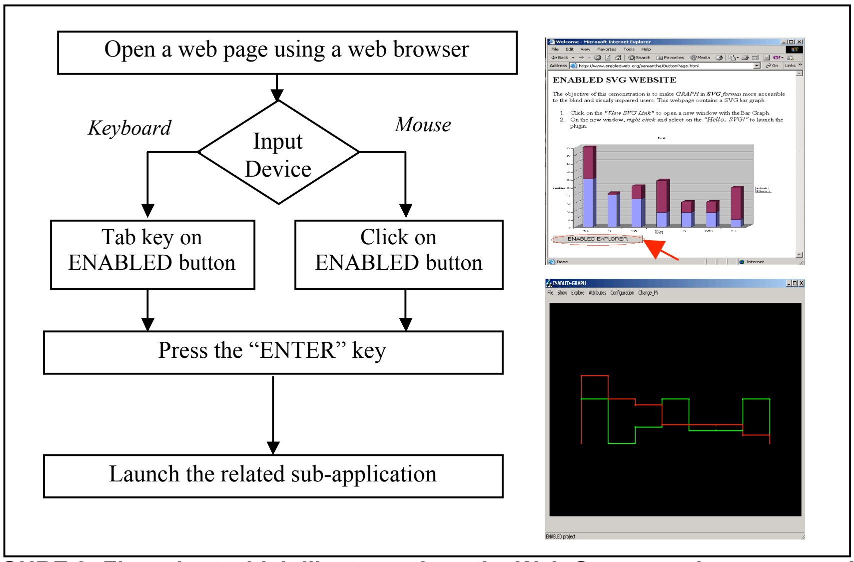


FIGURE 3: Flow chart which illustrates how the Web Content sub-system works

3.3 An example of a sub-application integrating with Web Content sub-system

The sub-applications are those small stand-alone applications that are responsible for the haptic-audio interaction and the data which has been extracted from the data file. It will then create an end interface to the user according to the assistive technologies available to the user and the user preferences. As applications can be developed independently, the adaptive system is also extensible, where any researchers can develop a new application to allow access to any kind of graphics by exploiting any types of assistive technologies. To illustrate how the Web Content sub-system works with sub-applications, a sub-application which enables the user to explore graphs in Excel format and SVG format by using the Phantom Desktop from SensAble Technologies [9], has been taken as an example in this paper to integrate with this adaptive system.

4. ADVANTAGES OF THE ADAPTIVE SYSTEM

This adaptive system is designed to investigate the feasibility of developing a system that is able to adapt to graphical Web content, assistive technologies and user preferences. As this system is compatible with both Internet Explorer and Mozilla Firefox browsers, users who are familiar with the use of these browsers, are able to utilize this adaptive system without having the inconvenience of learning a new browser. Furthermore, the adaptive system is designed according to how blind users use computers and Internet technology. As this application is developed based on the MSAA technology, it will not cause problems for screen readers to work smoothly with this system. Additionally, the involvement of users is kept to a minimum.

The potential impacts of this adaptive system will be beneficial to people from various realms. The most noticeable merit is to the end-users, especially the blind and visually impaired people who would like to know more about a Web graphic. Besides this, the adaptive architecture can act as a reference for other researchers or developers when developing an analogous architecture of their own. In addition to the objectives of developing an adaptive and interoperable architecture, this system is designed in a way that it is extensible and ready for the integration of future assistive technologies. Researchers or developers would be able to integrate their new stand-alone applications easily to this adaptive system, and it is foreseen that this adaptive system will become a coherent system which contains numerous sub-applications that are able to provide a good interface to its end user according to the user profile and preferences.

5. CONCLUSION & FUTURE WORKS

Currently, there are many ongoing research projects which aim to enhance Internet accessibility for visually impaired users. However, there is a need to have an adaptive system that is capable of handling various assistive technologies and allows them to interact between each other. This research intends to cater for user's preferences while providing an optimum interface for the users with visual disabilities via various types of assistive technologies available, to access the diverse formats of graphical web content. The sub-system responsible to detect the Web content and to launch the sub-applications has been developed and described in this paper. This adaptive architecture is believed to be able to bring massive improvements to the accessibility of visually impaired people when using the Internet. Evaluation is planned to be carried out to ascertain if the developed interface is accessible to visually impaired people and to gather information that is useful in the next phase of development. Further development on the architecture will be carried out to produce an adaptive, interoperable and extensible system.

ACKNOWLEDGEMENTS

This work is supported by ENABLED – an Integrated Project funded by the EU Framework 6 Programme, IST (FP6-2003-IST-2-004778), www.enabledweb.org. We would like to thank Teresa Gutierrez and colleagues from LABEIN who contribute to the main part of the development of graphs sub-application.

REFERENCES

[1] King, A., Evans, G. and Blenkhom, P. (2004) Webbie: a web browser for visually impaired people. Proceedings of 2nd Cambridge Workshop on UNIVERSAL ACCESS and ASSISTIVE TECHNOLOGY (CWUAAT), Cambridge.
[2] Craven, J. (2003) Access to electronic resources by visually impaired people. Information Research, Vol. 8, No. 4, pp. 156.
[3] O'Malley, M. and Hughes, S. (2003) "Simplified Authoring of 3D Haptic Content for the World Wide Web." 11th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, pp. 428-429.
[4] Way, T. and Barner, K. (1997) Automatic Visual to Tactile Translation, Part I & Part II: Human Factors, Access Methods and Image Manipulcation. IEEE transactions on Rehabilitation Engineering.
[5] Kawamura, H., Hakkinen, M. and DeMeglio, M. (2003) AMIS: the open source adaptive multimedia information system. Proceedings of 18th CSUN 2003 Conference on Technology and Persons with Disabilities, Los Angeles.
[6] Stephanidis, C., Paramythis, A., Sfyrakis, M., Stergiou, A., Maou, N., Leventis, A., Paparoulis, G. and Karagiannidis, C. (1998) Adaptable and adaptive user interfaces for disabled users in the AVANTI project. Proceedings of the 5th International Conference on Intelligence in Services and Networks (IS&N'98), Belgium.
[7] Holland, S., Ester, M. and Kiessling, W. (2003) Preference mining: a novel approach on mining user preferences for personalized applications. Proceedings of 7th European Conference on Principles and Practice of Knowlegde Discovery in Databases (PKDD '2003), Dubrovnik, Croatia.
[8] Penney and Associates, Ltd. (1996) Internet access for people with disabilities. Retrieved from: http://www.digitalgovernment.org/library/library/html/mwe_internetaccess_peoplewdisabilities_1996.jsp.
[9] Sensable Technologies website: www.sensable.com/

Technologies for Emotion-Aware Systems

Christian Peter, Steffen Mader, Karina Oertel, Michael Blech, Randolf Schultz, Jörg Voskamp, Bodo Urban
Fraunhofer Institute for Computer Graphics, J. Jungius Str. 11, 18059 Rostock, Germany
{christian.peter, steffen.mader, karina.oertel, michael.blech, randolf.schultz, joerg.voskamp, bodo.urban}@igd-r.fraunhofer.de

Emotions influence physiological processes in humans and are controlled by the autonomous nervous system. Emotions result in generally observable variations in a human's behaviour or physiological parameters that can be accessed with modern technology. For use in HCI contexts, emotion detection technology has to be minimally obtrusive and needs to be easy to use. In this article we present the latest developments in the field of emotion recognition technology undertaken at Fraunhofer IGD, Rostock, with focus on easily measurable physiological parameters. Methods of data mining and knowledge discovery used to build emotion classifiers are briefly described as is the underlying communication framework for networking the emotion processing components involved. Finally, a novel, configurable emotion induction application is introduced along with an emotion visualisation tool, which have been designed especially for emotion studies in the HCI context.

Emotion, emotion detection, emotion sensors, affective technology, emotion induction

1. INTRODUCTION

Emotions manifest themselves in physiological changes controlled by the autonomous nervous system [15, 22]. There are several physiological parameters which can be generally observed by humans, such as facial features, gestures, changes of the voice, or the breath rate. Other parameters are less obvious, like heart rate, blood pressure, skin temperature, electro-dermal activity, or pupil dilation.

Sensing and recognizing emotions with computers is a challenging undertaking, requiring the integration of hardware (sensors), mathematical methods for data enhancement and filtering, and pattern recognition and classification. While facial expressions are one of the most obvious manifestations of emotions [9], detecting them is still a challenge [6] although some progress has been made in the recent years (cf. [1, 10]). Problems here arise especially when the observed person moves about freely. Speech parameters have also been examined for correlations with emotions, with increasing acceptable results [5, 7, 14, 16]. Challenges here are similar to those of facial features. As the person moves about, it is necessary to track her position to get adequate measures of the speech signal. In contrast to facial features, speech delivers emotion information on several levels, or channels, which can be analysed independently from each other. So it is possible for example to analyse the content of utterances for emotive words or phrases, even when the speech signal is not of sufficient quality for signal analysis [27]. Gesture and body movement/posture also contain indicators of emotions. These are less researched than speech or facial features, but entail similar difficulties with moving subjects. There are two options to infer emotions in gesture: symbols (e.g. thumbs up), and semantics based on the fact that distinct emotions are associated with distinct qualities of body movement, such as tempo or force. Emotion-related changes of physiological parameters have been studied for a long time [e.g. 3, 8, 13, 20] and can be considered to be those emotional signs best understood today. A number of proof-of-possibility studies for emotion detection in non-lab settings have been performed, using either commercially available sensor equipment for physiology data or experimental devices. Among the latter are stationary, mobile, or wearable systems [12, 21, 23, 26, 28], or furniture equipped with sensors [2, 29].

As one can see, unobtrusive emotion detection in real-world applications poses several challenges to technology. At the department for Human-Centered Interaction Technologies of IGD Rostock we focus on developing technology for human-computer interaction for use in everyday settings. We soon realised the shortcomings of existing technology, which has been designed either for medical applications, psychology studies, or biofeedback applications. In each case, people act as “subjects” and are expected to accept electrodes and wires attached to them and to agree to being monitored. This is in contrast to being a “user.” To develop robust, unobtrusive realworld emotion aware systems we identified the following factors necessary to work on: a variety of complementary sensing devices that are robust, reliable, non-intrusive, and easy to use; a suitable way to communicate and integrate emotion-related data from all sorts of data sources; powerful algorithms for real-time analysis an classification of emotion-related data; and, for emotion studies, unsuspicious means to induce emotions.

The technology demonstrator described here presents the latest developments from IGD in unobtrusive detection of signs of emotion in users of computer systems. The focus is on measuring physiological parameters like heart rate, skin temperature, skin conductivity, and speech signal analysis. The demonstrator is also equipped with a configurable emotion induction application based on the Tetris game. Presented in this paper is the underlying communication framework for emotion processing components.

2. EMOTION SENSING

2.1 Physiology Sensors

Currently available sensor systems such as Thought Technologies' Procomp family, Mindmedia's Nexus device, or BodyMedia's SenseWear system are widely used in emotion research besides traditional medical devices like electrocardiographs (ECG), electroencephalographs (EEG), and electromyographs (EMG) for collecting emotionrelated physiological data. Apart from the SenseWear system, they use traditional electrodes as sensor elements. These are attached to the subjects with tape or Velcro fastener, the wires being directly connected to the data collecting device. This not only irritates and distracts users from their task, but also hinders free and natural movements [12, 21]. The collected data is either stored locally on the device, or transmitted directly to the processing computer. In either case, the data can only be accessed, viewed and analysed using the manufacturer's software and is not available instantly to other applications.

Affective applications, however, need direct and immediate access to the data to allow for continuous adaptation of the system to an ever-changing user state. Also, there should always be sensible data available, freeing the programmer from caring about lost connections, transmission errors, badly fitted electrodes, and other technical side aspects. For the same reason, data should be made available to processing applications in engineering units, avoiding inclusion of sensor-specific formulae into applications and eliminating the risk of conversion errors.

FIGURE 1. The (D)EmoKoffer technology demonstrator. The notebook computer features the EmoTetris emotion induction application, with emotion visualisation in the bottom right corner. The EREC emotion recognition system can be seen right to the computer, consisting of a sensor glove, a chest belt, and a data collection unit. The microphone for speech recording is integrated in the computer and can not be seen.

At Fraunhofer IGD we have developed a wireless and easy to use sensor system for collecting emotion-related physiological parameters skin resistance, skin temperature, and heart rate. The system has been designed for emotion researchers who want to examine emotional issues outside of laboratory contexts; for software developers who want to make use of emotion information in their systems;; and for psychologists who want to perform their studies in a natural setting without the irritating and distracting effects of wires on their subjects. The system has been designed for use by lay-persons in an everyday environment, mobile or stationary, and without putting restrictions on the user's behaviour. It gives the researcher and programmer more freedom in handling the data, providing the measurements conveniently in engineering units. Developed with the researcher and application developer in mind, the device has robust and reliable error handling and diagnosis mechanisms, guaranteeing sensible data continuously being available along with reliability information. The system is small, light-weight, functions wirelessly, transmits data immediately and is able to store data locally. It operates for several days on one battery pack. It has an open architecture and sends out the data in an open format, allowing software developers to easily incorporate the device into their systems as emotion sensing input source.

2.2 Emotion Detection in Speech

A human's voice reveals a lot on the speaker's emotions. In many contexts, this additional information sometimes is more important than the spoken words. It provides additional information about the meaning of a message, the satisfaction or anger of the speaker, for example if, urgent action or assistance is needed. At Fraunhofer IGD we work on tools to extract this additional information from a user's speech. Our research covers quality features of the speech signal, prosody, and semantic hints. Apart from common methods to extract emotion-related features of the speech signal [cf.11, 17], we are developing new algorithms for real-time identification of emotional states.

The recognition tool in this demonstrator utilizes signal analysis only and allows to distinguish between the emotional states angry/furious, fearful, disgusted, sad, and bored with an accuracy of 75 to 85%, which is close to that of human listeners. These results were obtained on the data from the speech corpus of the Technical University of Berlin, which contains short sentences spoken by actors.

Algorithms applied to short utterances from an online banking system allow us to assess the users' state concerning arousal and valence (pleasantness) with an accuracy of up to 70%. Developing these classifiers has been more challenging, because the speech examples are from natural situations and thus more ambiguous and unclear concerning the underlying emotion. Additional filter operations had to be implemented because of the partially bad signal quality of the recordings. Current developments promise to improve these results by exploiting prosodic and semantic information.

3. KNOWLEDGE DISCOVERY AND DATA MINING

Sensing emotion-related information delivers a huge amount of data with an enormous quantity of parameters being extracted from them [cf. 4]. From the physiology data heart rate, skin resistance and skin temperature, for instance, 22 emotion-relevant attributes are currently selected. The speech signal yields 1200 attributes potentially relevant for emotion analysis with more than 60 usually being exploited.

Datamining techniques are best suited since they allow to deal with big data sets and to examine them without previous specification of hypotheses or parameters to use. The extracted information can then be used to define attributes, statistical methods, learn algorithms and classification concepts for integration into emotion detection classifiers, cf. [24]. The filter operations and classifiers used in this technology demonstrator have been implemented using the OmniRoute framework described in the following section. OmniRoute allowed us to process and classify the received physiological data in real-time, for instance for online visualisation of the classification results.

4. COMMUNICATING EMOTION INFORMATION

Affective computing systems consist of a number of dedicated components which, in most of the cases, were developed as independent stand-alone applications. Those components usually are designed around the task they are targeting. Decisions regarding the platform, the programming language to use, and the ports / interfaces to be provided are made mainly when the targeted problem in focus. Other considerations, such as networking capabilities or compatibility with systems or applications not directly linked with the particular task were laid aside or neglected at all.

Integrating such components into a complex system is a challenging task. The components provide a limited number of different interfaces, each component delivers or expects data in dedicated native formats, and there is a huge variety in the data concerning their complexity, continuity, reliability, size, and transmission speed.

To overcome these challenges we have developed OmniRoute, a framework which can be considered as a system construction kit, where diverse components are wrapped into smoothly fitting building blocks that can be freely arranged and allow for a wide range of experiments [18]. Our framework provides the infrastructure for connecting the components, an easy to use configuration mechanism, and unified data handling schemes suitable for realtime use. Integration is accomplished by adaptors for each component, which provide the required uniformity in data representation. Moreover, the adaptors are designed to allow handling of different components in a similar manner. In their entirety, these adaptors form an additional abstraction layer between the data sources and processing applications.

5. EMOTION INDUCTION AND VISUALISATION

Emotion induction for emotion studies in the HCI context is an intricate endeavour. The experiment application should fulfil the following requirements: Firstly, as most users want to reach a goal when interacting with a computer (like writing an article or buying a ticket), the task should be of some relevance to the subject while in the test. Secondly, for smooth transition from one emotion to another it should be possible to change the application's procedure without causing suspicion in the user. For example, it should be possible to induce contentment first, then turn to happiness and proceed with boredom before putting the user in a helpless and later furious state.

Our chosen emotion induction tool, EmoTetris, meets these requirements [19]. It is based on the Tetris gamewhich is known for its simple but intriguing gameplay. Games in principle affect the mood of the player [25] andTetris' popularity should make sure that most subjects will like the game and are trying to be successful in playingit. Since the task of the player is to react on different items that occur randomly, the experimenter has a high levelof control. For example, by varying the speed or form of the bricks, different emotions can be induced. We also introduced additional features, good and bad wizards that give or take bonus points, and special events with audible and visual effects motivating or irritating the user. EmoTetris has proved to be the most popular and hence most successful emotion induction tool at Fraunhofer IGD.

For verifying the quality of our emotion recognition results we developed various means to visualize the classified emotions. Figure 1 shows a visualisation realised as a plug-in within EmoTetris which is used to monitor ongoing emotions on-line. The visualisation contains: (a) a star plot for showing the variance of basic emotions in the dimensions of valence and arousal plus current value; (b) verbal presentation of recognized emotions: and (c) a comic face for an efficient representation of predominant emotion.

REFERENCES

[1] Aleksic P.S., Katsaggelos A. K. (2005) - Automatic Facial Expression Recognition Using Facial Animation Parameters And Multi-Stream Hmms. IEEE Trans. on Sig. Proc. Supplement on Secure Media, 2005.
[2] Anttonen, J. & Surakka, V. (2005). Emotions and Heart Rate while Sitting on a Chair. In CHI ’05 conference proceedings, pp. 491-499. New York: ACM Press.
[3] Ax, A., 1953. The physiological differentiation between fear and anger in humans. In Psychosomatic Medicine 55 (5), pp. 433-442. The American Psychosomatic Society.
[4] Blech M., Peter C., Stahl R., Voskamp J., Urban B. (2005). Setting up a multimodal database for multi-study emotion research in HCI. Proceedings of the HCI International Conference, Las Vegas, 2005.
[5] Burkhardt, F. & Sendlmeier, W. F., 2000. Verification of Acoustical Correlates of Emotional Speech using Formant-Synthesis. In ISCA Workshop on Speech and Emotion.
[6] Cowie, R., Douglas-Cowie E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., & Taylor, J.G. (2001).Emotion recognition in human computer interfaces. IEEE Signal ProcessingMagazine, January, 2001.
[7] Douglas-Cowie, E. , Cowie, R., Campbell, N. (2003). Speech and Emotion (2003). Speech Communication 40 (2003), 1-3. Elsevier Science.
[8] Ekman, P., Levenson, R. W. & Friesen, W., 1983. Autonomic Nervous System Activity Distinguishes among Emotions. In Science 221. The American Association for Advancement of Science.
[9] Ekman P., Davidson R. J. (Eds.) (1994). The Nature of Emotion: Fundamental Questions. Oxford University Press, New York.
[10] Fasel B., Luettin J. (2003). Automatic Facial Expression Analysis - A Survey. Pattern Recognition, 36, 259- 275.
[11] Fernandez, R. (2004). A Computational Model for the Automatic Recognition of Affect in Speech. Ph.D. thesis, Massachusetts Institute of Technology.
[12] Haag A., Goronzy S., Schaich P., Williams J. (2004). Emotion Recognition Using Bio-sensors: First Steps towards an Automatic System. In André et al (Eds.): Affective Dialogue Systems, LNCS 3068, Springer- Verlag, pp. 36-48.
[13] Herbon A., Peter C., Markert L., van der Meer E., Voskamp J. (2005). Emotion Studies in HCI – a New Approach. Proceedings of the 2005 HCI International Conference, Las Vegas, 2005.
[14] Küstner D., Tato R., Kemp T., Meffert B., 2004. Towards Real Life Applications in Emotion Recognition; Comparing Different Databases, Feature Sets, and Reinforcement Methods for Recognizing Emotions from Speech. In André et al (Eds.): Affective Dialogue Systems, LNCS 3068, Springer-Verlag.
[15] Lane, R.D., Nadel L. (Eds.) (2000). Cognitive Neuroscience of Emotion. Oxford University Press.

[16] Laukka P., Juslin P. N., Bresin R., 2005. A dimensional approach to vocal expression of emotion. Cognition and Emotion 2005, 19 (5), Psychology Press, 633-653.
[17] Lee, C. M. and Narayanan, S. S. (2005). Toward detecting emotions in spoken dialogs. IEEE Transactions on Speech and Audio Processing, 13(2).
[18] Mader, S., Peter, C., Göcke, R., Schultz, R., Voskamp, J., Urban, B. (2004): A Freely Configurable, Multimodal Sensor System for Affective Computing. In: André et al. (Eds.): Affective Dialogue Systems, LNCS 3068, Springer-Verlag Berlin, Heidelberg, New York, 313-318.
[19] Oertel, Karina; Schultz, Randolf; Blech, Michael; Herbort, Oliver; Voskamp, Jörg; Urban, Bodo:EmoTetris for Recognition of Affective States. Proceedings of the 2005 HCI International Conference, Las Vegas, 2005.
[20] Palomba, D. & Stegagno, L., 1993. Physiology, Perceived Emotion and Memory: Responding to Film Sequences. In Birbaumer, N. & Öhman, A. (Eds.): The Structure of Emotion, pp. 158-168. Toronto: Hogrefe & Huber Publishers.
[21] Peter, C., Ebert E., Beikirch, H. (2005). A Wearable Multi-Sensor System for Mobile Acquisition of Emotion- Related Physiological Data. Proceedings of the 1st International Conference on Affective Computing and Intelligent Interaction, Beijing 2005. Springer Verlag Berlin, Heidelberg, New York.
[22] Picard, R.W. (1997). Affective Computing. M.I.T. Press, Cambridge, MA.
[23] Picard, R. W, and Healey, J., (1997). Affective Wearables, Personal Technologies Vol 1, No. 4 , 231-240. [24] Picard R.W., Vyzas E., Healey J. (2001). Toward Machine Emotional Intelligence - Analysis of Affective Physiological State. IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol 23 No. 10 October 2001.
[25] van Reekum, C.M., Johnston, T., Banse, R., Etter, A., Wehrle, T., & Scherer, K.R. (2004). Psychophysiological responses to appraisal dimensions in a computer game. Cognition & Emotion, 18(5), 663-688.
[26] Scheirer, J., Fernandez, R., Klein, J., Picard, R.W., 2002. Frustrating the user on purpose: a step toward building an affective computer. Interacting with Computers 14 (2), 93–118.
[27] Schröder, M., 2000. Experimental Study of Affect Bursts. In ISCA Workshop on Speech and Emotion.
[28] Sung M., Pentland A., 2005. Minimally-Invasive Physiological Sensing for Human-Aware Interfaces. Proceedings of the 2005 HCI International conference.
[29] Tan H. Z., Slivovsky L.A., Pentland A. (2001). A Sensing Chair Using Pressure Distribution Sensors. IEEE/ASME Transactions On Mechatronics, Vol. 6, No. 3, September 2001.

Task Modeler: Innovative Tooling for Established Methods

Paul Englefield, IBM Global Business Services, Mark Farmer, IBM Software Group, Nik Mottershead IBM Global Business Services, Mark Tibbits, IBM Software Group, Ian Wells IBM Software Group
PO Box 41 Birmingham Road, Warwick, CV34, 5JL, United Kingdom, 01926-464329
paul_englefield@uk.ibm.com

Task Modeler provides effective tools support for User Experience practitioners who need to work with design stakeholders to create, analyse, and share task-based models of human activity. The tool is practical, methods-based, analytical and extensible. It is currently used by design practitioners in IBM. A technology preview can be downloaded from the IBM alphaWorks site.

Task analysis, task modelling, usability, design, User Engineering, DITA, practitioner support, tools support

1. INTRODUCTION

Task Modeler is an Eclipse-based [1] software tool intended to support practitioners in applying established design methods through robust tooling. Task Modeler facilitates the construction of a family of "task-based" models that complement a core hierarchical decomposition of tasks with a rich vocabulary of related elements, properties and heterarchical relationships. Task Modeler delivers a method as a "dialect". Each such dialect defines the elements required to create models appropriate to that method. In practice, dialects can also be constructed to support methods that are not inherently task-based.

Tools such as CTTE [2] and Tarmot [3] embody methods innovation such as ConcurTaskTrees and Diane+. Task Modeler, however, emphasizes new approaches to supporting practitioners in putting methods to work. Innovative aspects include: a) architectural separation of a modelling engine from "pluggable" dialects to enable convenient integration of new and refined methods; b) a practitioner-friendly user experience for rapidly creating, managing and navigating large models; c) dynamic visualisation of property values and heterarchical relationships.

The exhibit should be of interest to practitioners with an interest in robust and flexible tooling to support practical application of task analysis methods to both problem space analysis (e.g. [4]) and task-based abstract design (e.g. [5]).

2. NODE-TYPES, PROPERTIES, AND RELATIONSHIPS

The content of a dialect is defined as a set of node-types, properties, and relationships. For example, as well as *Task* nodes required by a classic HTA (Hierarchical Task Analysis) model [6], the UE (User Engineering) [7] dialect provides additional node-types to model contextual elements such as *Role*s and *Environment*s connected by typed heterarchical relationships such as *works-with* and *related-to*. Each node-type is enriched by a specific property set. For example, the properties associated with a *Role* node are equivalent to a user profile. Properties can include strings, numbers, ratings, file references, web references, and keywords. Keywords can have controlled or extensible vocabularies as required. For example, the keyword property *cognitive factor* has a pre-defined controlled vocabulary including *attention*, *memory*, and *perception*, whereas the keyword property job title is extensible to enable a researcher to develop a vocabulary of titles.

3. PRACTICALITY

In commercial design practice, modelling is frequently carried out in the field or in a collaborative design workshop. For example, documenting current user behaviour may require concurrent observation, interviewing, and modelling. Likewise, exploring alternative design responses involves rapidly transcribing and reorganizing design proposals from collaborators. In both cases, responsiveness is critical [8] to avoid "transcription lag". Models can also rapidly become inconveniently large; finding and moving to a specific section can be a challenge in a "live" modelling situation.

Task Modeler uses a number of user interface techniques to ensure that the interface is sufficiently responsive and flexible to enable the scribe to "keep up" with the process. These include: extensive keyboard support; alternative navigation schemes; multiple zoom levels; selective data hiding; and a choice of presentation styles. Because tools are rarely used alone, Task Modeler provides flexible imports and exports to enable interoperability with office applications and other design tools. To encourage best practice, each dialect is supported by a set of teaching and coaching resources. These include a method guide, a node-type guide, and a set of validators to offer interactive critiques of work in progress.

4. METHODS-BASED MODELLING

Task Modeler can be considered as a tailored delivery mechanism for methods. To create a new model, a user initially selects a dialect. Task-based dialects include: HTA; UE Roles and Goals; and DITA (Darwin Information Typing Architecture) [9]. Task Modeler is also architected to support a broader family of methods. Candidate dialects might include: QOC (Questions, Options, Criteria) [10], IA (Information Architecture) (e.g. [11]), and Value-based design. Each dialect defines a method-specific modelling experience in terms of the following elements: an ontology that specifies the required node-types, properties, controlled vocabularies, and heterarchical relationships; dialect-specific imports and exports; dialect-specific reports, dialect-specific validators; and dialect-specific guidance. Table 1 shows illustrative examples of these elements for a hypothetical QOC dialect.

Node-type	Question, Option, Criterion, Argument
Properties (for Criterion)	Name, Description, Sponsor, Importance, …
Controlled vocabulary (for Sponsor)	Business, Project, Technical, User, …
Relationships (for Criterion)	Supports [Argument] Related to [Criterion]
Reports	QOC matrix Recommendations
Exports	Data to feed a design briefing service delivered as an RSS feed.
Validators	Checks for well formed QOC model with balanced, complete argumentation,

TABLE 1: Example dialect elements

5. PRACTITIONER CUSTOMIZATION

Although dialects are designed as integrated packages by subject matter experts, individual practitioners may also need to adapt and extend the supplied methods. They can do so in various ways: by defining additional properties for a node-type; by defining additional values for a keyword property; and by “plugging in” user-developed export transforms. For example, while analysing the problem space for a Defence application, a practitioner might customize the UE roles and goals dialect by defining an additional property, *Rank*, for the *Role* node-type and specifying keywords such as Sergeant and Captain. By designing a custom export transform she could then transfer these extended role definitions into a document database such as Lotus Notes to enable a review by an extended team.

6. VISUALIZATION

Effective design requires both action and reflection documentation [12]. To support “conversation with materials” practice, Task Modeler uses colour, text, and symbols to visualize properties in order to identify regularities and hot spots. A practitioner can select any numeric or keyword property in order to visualize the distribution of that property’s values across the model. For example, a risk analyst might apply colour to highlight all tasks flagged as Safety-critical or Business-critical. Likewise a designer using the object-oriented OVID method [13] might apply an overlay of symbols to show the distribution of user-perceived objects. Similarly, a practitioner can select a non-hierarchical relationship to visualize. For example, a designer of collaboration solutions might visualize a works-with relationship as a network overlaid on a roles and goals decomposition.

7. EXAMPLE

Figure 1 shows a screen used by a practitioner analysing a Roles and Goals model for a coffee bar. The model shows two user roles (Barista and Cashier) taking and fulfilling beverage orders. These users work in a context (Coffee lounge) for an organization (Caterer). Each role has a goal supported by a set of tasks. For example, the Barista role has a goal of serving customers and tasks of taking, preparing, and handing over the order. The user experience designer has used the Property Visualization view to highlight safety-critical tasks (shown with a blue background) and to mark the task type (symbols). He has also applied a network visualization to show that the Barista works with the Cashier. Visualisation enables the model to "talk back" to the designer.

8. APPLICATIONS IN IBM

Task Modeler is used by three distinct communities at IBM: Information architects use the tool to design the DITA maps that define the high level structure of the documentation for IBM software products; User Engineering practitioners use it to analyse the roles, goals, and task structures that shape the design of user interfaces for hardware and software products; finally, user experience consultants use it to analyse problem spaces and design solutions within commercial consultancy projects. In all these cases, the motivation is to encourage rigorous analysis and design by delivering methods knowledge through productive, flexible tooling.

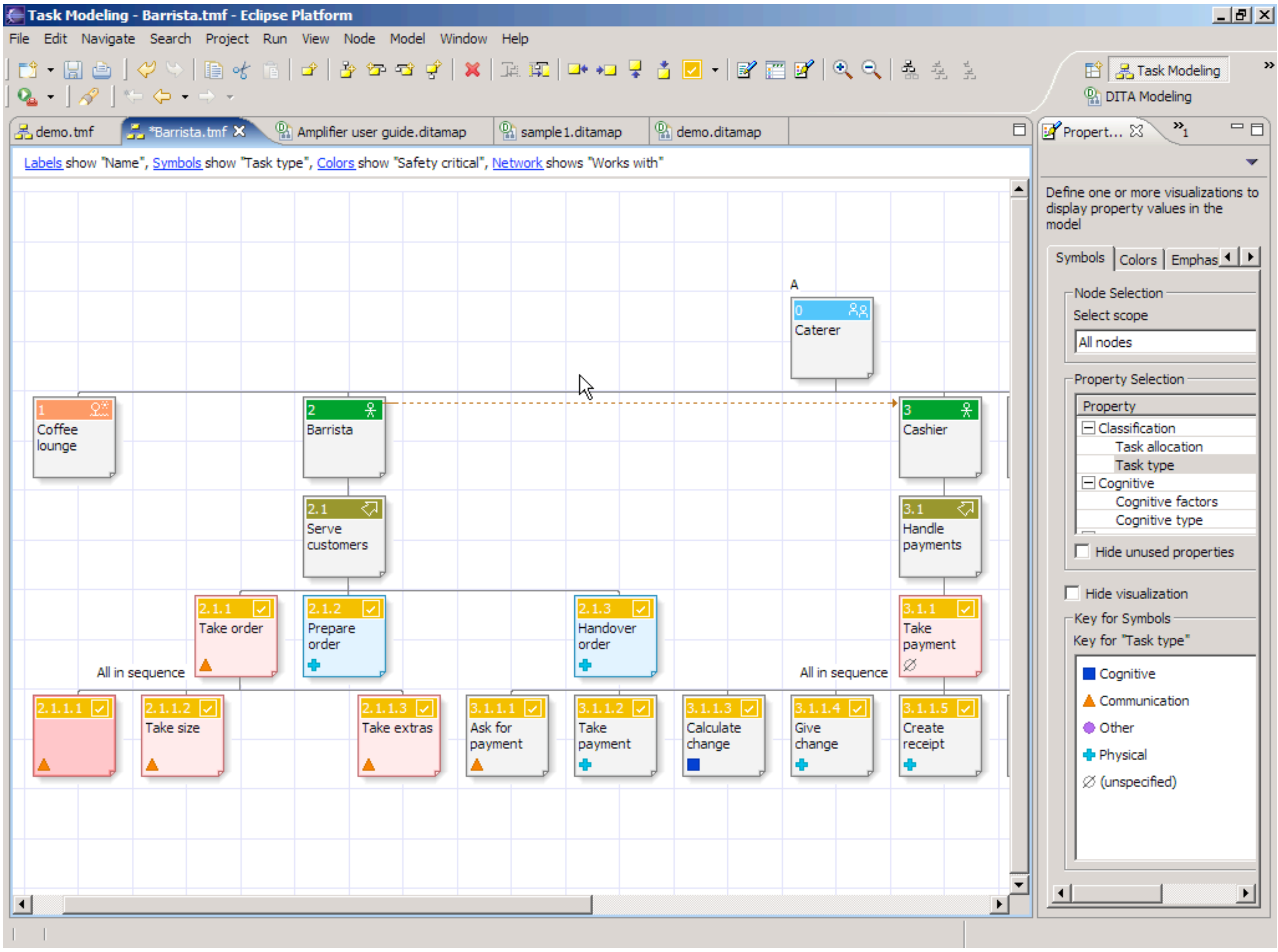


FIGURE 1: Screen shot: analysing user-perceived objects

9. AVAILABILITY

You can download a preview version of Task Modeler from alphaWorks [14]. *www.alphaworks.ibm.com/tech/taskmodeler.* alphaWorks is intended as a showcase for emerging technologies. It enables customers and researchers to get early insight into new technical thinking while offering IBM innovators an opportunity to gather feedback from potential users.

10. ABOUT THE INTERACTIVE EXPERIENCE

The design lead for Task Modeler will demonstrate the features described above and invite delegates to use the tool themselves to create a range of models. Feedback will be sought on the tool, the supported methods, and additional potential applications.

REFERENCES

[1] Eclipse *www.eclipse.org*

[2] Paris, C., Lu, S., Vander Linden, K. (2004) Environments for the Construction and Use of Task Models. In Diaper, D. and Stanton, N. (eds) *The handbook of Task Analysis for Human-Computer Interaction.* Lawrence Erlbaum Associates, New Jersey
[3] Paterno, F., (2004) An engineered notation for task models. In Diaper, D. and Stanton, N. (eds) *The handbook of Task Analysis for Human-Computer Interaction.* Lawrence Erlbaum Associates, New Jersey
[4] Hackos, J. and Redish, G. (1998) *User and Task Analysis for Interface Design*. John Wiley & Sons Inc, New York
[5] Newman, W. and Lamming, M. (1995) *Interactive system design*. Addison Wesley, Harlow
[6] Annet, J. and Duncan, K. (1967) Task Analysis and Training Design. *Occupational Psychology*, 41, 211-221
[7] IBM User Engineering *www.ibm.com/easy*
[8] Lambell, N.J., Ball, L.J., & Ormerod, T.C. (2000). The evaluation of Desperado: A computerised tool to aid design reuse. In R. Winder & S. McDonald (eds.) *People and computers XVI: 'Usability or else!'* Cambridge University Press , Cambridge, UK
[9] DITA *www.oasis-open.org/specs*
[10] MacLean, A., Young, R. Bellotti, V. and Moran T. (1991) Questions, options and criteria: Elements of design space analysis. *Human-Computer Interaction*, 6, 201-250
[11] Rosenfeld, L. and Morville, P.(2002) *Information Architecture for the World Wide Web*. O'Reilly, Sebastopol
[12] Schon, D. (1996) Reflective conversation with materials In T.Winograd (ed) *Bringing design to software*, ACM press, New York
[13] Berry, R. et al. (1998) *Designing for the user with OVID*. New Riders, Upper Saddle River, NJ
[14] IBM alphaWorks *www.ibm.com/alphaworks*

What is on the Backside of the Paper? From 2D Sketch to 3D Model

Day Chyi Ku, Sheng-Feng Qin, David K. Wright
School of Engineering and Design, Brunel University, Uxbridge, UB8 3PH, UK
DayChyi.Ku@brunel.ac.uk, Sheng.Feng.Qin@brunel.ac.uk, David.Wright@brunel.ac.uk

We demonstrate an intuitive and interactive sketching system that reconstructs 3D object incrementally from 2D freehand sketch. The system avoids conventional menu interface and provides interface mimicking pencil and paper that takes 2D freehand sketch as its direct input. The system offers a natural sketching interface by allowing sketching with discontinuous, overlapping and multiple strokes. The input sketch is a natural line drawing with hidden line removed that depicts a 3D object in real world. The system is configured to interpret sketch under isometric view, which is the most popular viewpoint among designers. The system automatically tidies-up the freehand 2D sketch and reconstructs it into a 3D object. The object can then be transformed, which provides immediate feedback to the designer. New sketch can be drawn over the 3D object, and then be reconstructed into 3D by referring to the existing 3D object from the current viewpoint. This method allows the designer adds new parts on the ``backside" of the initial sketch. The incremental modelling from sketch enables easy reconstruction of sophisticated 3D object comparing with the current 3D reconstruction algorithms. It also enables a more detailed object to be reconstructed.

3D modelling, freehand sketching, conceptual design, interactive sketching

1. INTRODUCTION

Freehand sketching is a fast and efficient way to visualise one's idea in conceptual design. A sketch is believed to assist designers by allowing their mental images to be expressed externally for further mental synthesis. With the introduction of faster computers, there is a stronger interest to implement a CAD system for conceptual design that can allow a designer to sketch a series of drawings and transfer them to 3D models directly and immediately. There are many advantages to have such a computerized system. Some of these relate to how the sketch data is stored (i.e., digital). For example, recent advancements in data storage make it relatively inexpensive to store many sketches in a single data drive (e.g., hard disk). Furthermore, with digital data, sharing and communication of ideas through sketches can be performed easily without lost of information (e.g., scanning sketches made on paper and converting them to image files decrease the quality of sketches and do not contain sketching sequence).

However, currently available commercial CAD systems, such as SolidEdge and Pro/ENGINEER, cannot create 3D object directly from freehand sketch. In particular, extensive menu selections are needed to create a 3D object. Such process is not as intuitive as sketching with pencil on paper, and hence is not suitable for conceptual design. CAD systems provide limited support for direct 3D reconstruction from freehand sketch because of the limitations of existing reconstruction algorithms. Most of the currently available 3D reconstruction approaches are susceptible to imperfection in line drawing such as in optimization-based approach [1, 2] and line labelling approach [3]. The robustness, i.e. the success rate to reconstruct a meaningful 3D object by the optimization-based approaches is dependent on the optimization method used. Furthermore, the optimization processing is computational expensive and often fails to converge to the optimal solution (if exists). Other approaches that involve perceptual observations and heuristic rules application [4-6] strictly limited the reconstructed models to symmetry, perpendicular and axis-aligned-objects.

Although there are studies on enabling direct 3D object reconstruction from freehand sketch [6-10], most of these proposed solutions are limited in providing a natural sketching interface. For example, these systems interpret the input sketch in a specific manner that some gestures actually mean the 3D reconstruction commands. As such, the development of a sketch system that provides direct 3D reconstruction from freehand sketch (e.g., CAD tools for conceptual design) remains a challenge that has yet to be fully addressed.

We propose an interactive freehand sketching system to assist the designer in the early designing stage. The system can provide an intuitive calligraphic interface that allows the designer to sketch out their desired shapes without enforcing gesture sketching, meanwhile avoids excessive menu selection that adds overhead to the design process. In addition, it provides a more natural and intuitive sketching environment by enabling discontinuous strokes and overtracing, and interpreting natural line drawing (hidden line removed sketch). The system also is able to interpret the freehand sketch and reconstruct it into 3D object in real-time. It is able to transform the 3D object for designer to evaluate their sketch from different viewpoint. Furthermore, it support non-photorealistic rendering (NPR) for consistent visual feedback. Finally, the system should be interactive so that the designer can work on his/her sketch progressively.

2. METHODOLOGY

The input to our system is the freehand sketch that is captured online through a calligraphic interface. Figure 1 shows the overall process in the system. We provide a natural and intuitive sketching environment by enabling multiple stroke sketching. The sketching movement is captured by recording the screen coordinates of the stylus over a fixed interval time while it is pressed and dragged on the screen.

The data is interpreted by a series of tidy-up processes to produce a vertex-edge graph. The process consists of four stages: stroke classification, strokes grouping and fitting, 2D tidy-up with endpoint clustering and parallelism correction, and in-context interpretation. Strokes are first classified into lines and curves by a linearity test. It is followed by an innovative strokes grouping process that handles lines and curves separately. The grouped strokes are fitted with 2D geometry and further tidied-up with endpoint clustering and parallelism correction. Finally, the in-context interpretation is applied to detect incorrect stroke interpretation based on geometry constraints and to suggest a most plausible correction based on the overall sketch context (more details are discussed in [11]). The graph consists of 2D coordinates representing connectivity of tidied-up line segments from sketched strokes. The system has the advantage of interpreting overtracing as the sketch content. It also allows modification and adjustment to be made by sketching over the existing sketch.

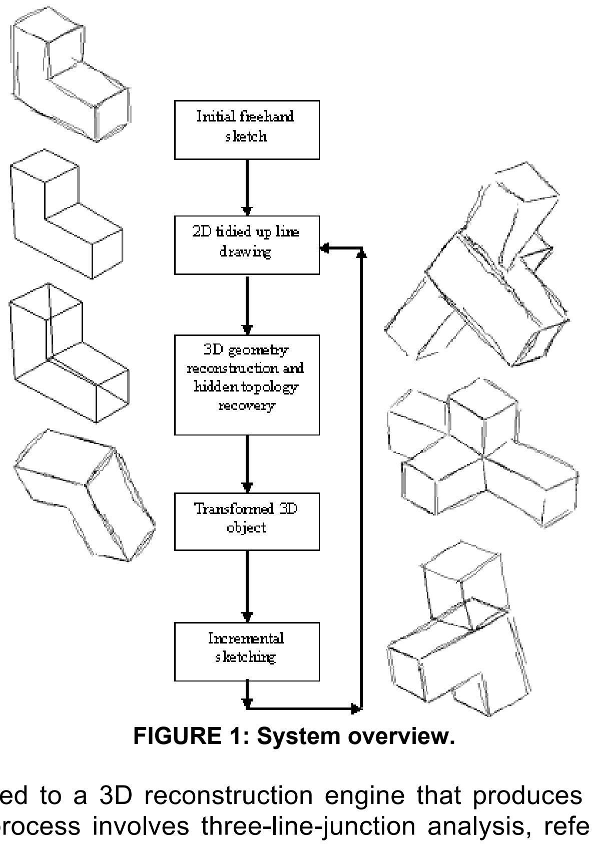


FIGURE 1: System overview.

After that, the graph is passed to a 3D reconstruction engine that produces a 3D object depicted by the sketch. The reconstruction process involves three-line-junction analysis, reference junction determination and vertex calculation, hidden topology recovery and planarity enforcement. The 3D object can then be transformed so that it can be viewed from different viewpoints with NPR appreance.

New sketch can be added directly onto the reconstructed 3D object in the current view. After the sketching session, 2D tidy-up of the new sketch is carried out. It is followed by second round 3D reconstruction process that reconstructs the new sketch into 3D object. The new 3D object is incorporated with the existing 3D object under the current viewpoint. After reconstruction, the new and existing objects will be treated as one object and can be transformed together. The designer can also sketch directly on the updated object after transformation. The sketching, interpretation and reconstruction process is repetitive and compatible to the conceptual design sketching activities. With the functionalities provided by our system, the designer is able to sketch on any part, even the "backside" of the object, enabling a sophisticated design to be presented in a single file.

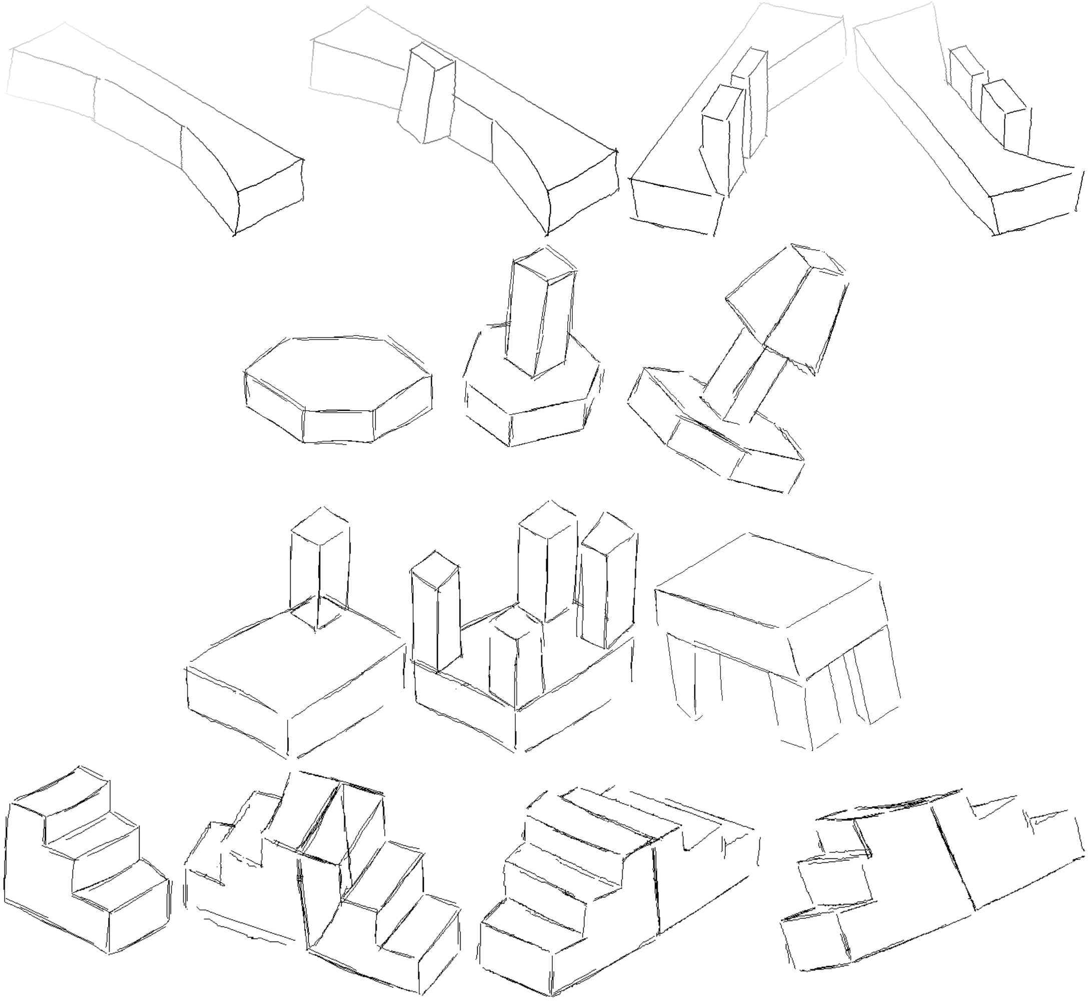


FIGURE 2: 3D objects reconstructed incrementally with our freehand sketching system.

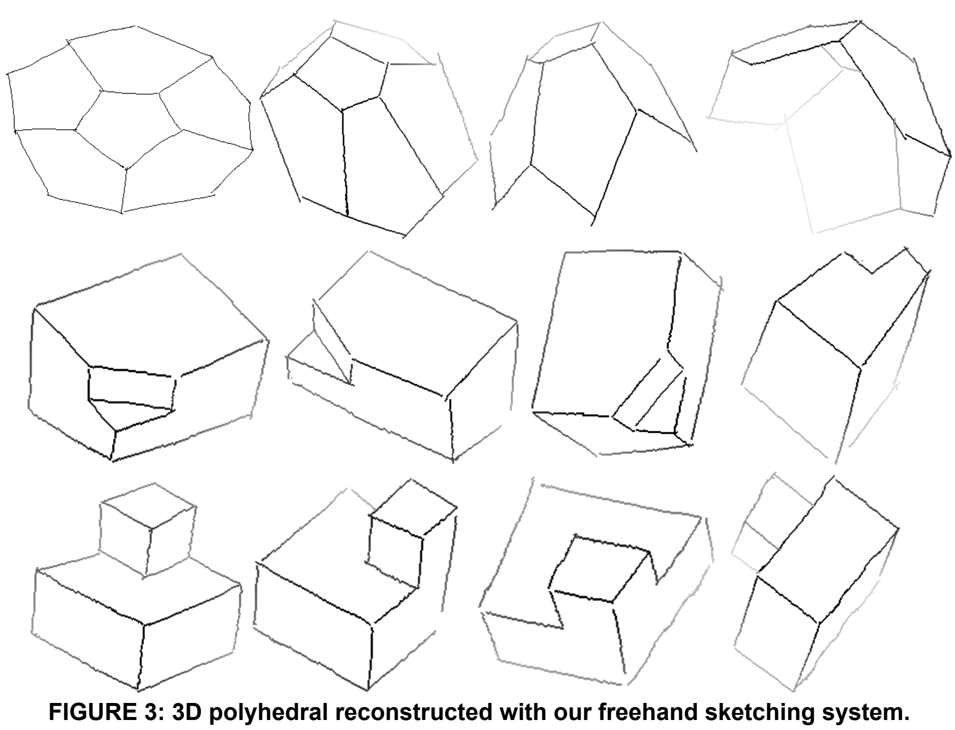


FIGURE 3: 3D polyhedral reconstructed with our freehand sketching system.

3. RESULT

Figure 2 shows examples of the progressive 3D reconstruction from incremental sketching with our sketching system. Our system is able to reconstruct non-axis-aligned, non-symmetry and non-perpendicular polyhedral (Figure 3). The objects are interpreted based on the freehand input without application of ambiguous heuristic rules e.g. parallel and perpendicular lines, and object symmetry. However, they are adjusted based on the unambiguous planarity geometry constraint. Meanwhile, the 3D objects are rendered with sketchy strokes, which make the object appeared similar to the input sketch from different viewpoint.

4. CONCLUSION AND FUTURE WORK

The intuitive and interactive sketching system discussed in this paper aims to support conceptual design sketching. The investigation and development of this system have made several contributions to the HCI and object modelling research. First, it supports natural freehand sketching with overtracing, which allows the designer to sketch more freely as sketching with pen and paper. Second, an innovative 3D reconstruction algorithm has been developed to interpret sketch into 3D object. The algorithm has the advantages of being simple, robust, insusceptible to input inaccuracies, thus suitable for online sketching application (with the potential to be used in Palmtop). Besides lifting the needs to enforce gesture sketching, the algorithm is able to reconstruct more general objects compared to the gesture-based system. It also allows designer to sketch directly on any part of the reconstructed 3D object, before or after transformation. This means new sketch can be added to any part of the reconstructed object and enables designer to sketch from viewpoints that are convenient to them. In addition, it can reconstruct the new sketch into 3D to be incorporated with the old 3D object. Finally, the rendering output is in non-photorealistic form (NPR) with the appearance similar to the original sketch. This is to ensure a consistent sketch style presentation on screen, which helps the designer to focus on the design instead of being distracted by the output appearance. The depth cue of the object is indicated by various grey tones, with darker tones indicating the closer end and lighter tones indicating the further end.

The system presented here is part of our final personalised-sketched-based 3D modelling and rendering system. Further study is required for implementation of simple curve objects reconstruction and non-photorealistic rendering. The reconstructed object's appearance can be improved by having shadows and textures on the surfaces of the object for more detailed sketch.

REFERENCES

[1] H. Lipson and M. Shpitalni, "Optimization-based reconstruction of a 3D object from a single freehand line drawing," CAD Computer Aided Design, vol. 28, pp. 651-663, 1996.
[2] P. Company, M. Contero, J. Conesa and A. Piquer, "An optimisation-based reconstruction engine for 3D modelling by sketching," Computers and Graphics (Pergamon), vol. 28, pp. 955-979, 2004.
[3] P. A. C. Varley, R. R. Martin and H. Suzuki, "Frontal geometry from sketches of engineering objects: is line labelling necessary?" Computer-Aided Design, vol. 37, pp. 1285-1307, 2005/10.
[4] D. Lamb and A. Bandopadhay, "Interpreting a 3D object from a rough 2D line drawing," in VIS '90: Proceedings of the 1st Conference on Visualization '90, 1990, pp. 59-66.
[5] M. Contero, F. Naya, J. A. Jorge and J. Conesa, "CIGRO: A minimal instruction set calligraphic interface for sketch-based modeling." in ICCSA (3), 2003, pp. 549-558.
[6] F. Naya, J. A. Jorge, J. Conesa, M. Contero and J. M. Gomis, "Direct modeling: From sketches to 3D models," in 1st Ibero-American Symposium on Computer Graphics - SIACG02, 2002,
[7] S. F. Qin, D. K. Wright and I. N. Jordanov, "From on-line sketching to 2D and 3D geometry: A system based on fuzzy knowledge," CAD Computer Aided Design, vol. 32, pp. 851-866, 2000.
[8] L. Eggli, C. Hsu, B. D. Bruderlin and G. Elber, "Inferring 3D models from freehand sketches and constraints," CAD Computer Aided Design, vol. 29, pp. 101-112, 1997.
[9] A. Shesh and B. Chen, "SMARTPAPER: An Interactive and User Friendly Sketching System," Comput. Graphics Forum, vol. 23, pp. 301-301, 2004.
[10] R. C. Zeleznik, K. P. Herndon and J. F. Hughes, "SKETCH: An interface for sketching 3D scenes," in SIGGRAPH '96: Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, 1996, pp. 163-170.
[11] Day Chyi Ku, SF Qin and David Wright, " Interpretation Of overtracing freehand sketching for geometric shapes," in Proceedings of WSCG’2006, 2006, pp. 263-270.

A Real-Time Spatial Measurement Interface for Emotional Evaluation of Temporal Media

James Ohene-Djan and Vladislav Sushko
Department of Computing, Goldsmiths College, University of London, New Cross, London, SE16 6NW
j.djan@gold.ac.uk and ma202vs@gold.ac.uk

In this paper we present a real-time advanced spatial measurement interface (RASMI), designed to allow users to engage with temporal media presentations, such as party political television broadcasts, in order to imprecise intuitive responses or emotions. The interface is based on an equilateral triangle that records user input as three sets of data points in real time over a two-dimensional space. This information is then used to calculate the levels of emotional uncertainty relative to other choices using a triangle-based spatial measurement technique.

Levels of emotional uncertainty are represented using a set of 2D Interactive graphs that link user responses to points in media playback timelines. 2D Interactive graphs not only provide a visual representation of the data flow but also act as an extremely engaging backtracking tool enabling easy identification of when within a media presentation and with what degree of uncertainty a user wished to register an emotion. Information captured is also visualized by means of an Equilateral Triangle Scatter Graph and line graphs.

The novelty of RASMI is its ability to allow groups of users to experience temporal media presentations while concurrently providing real-time data on their emotional experience. The RASMI is software runs on any PC, that's supports digital video, and a mouse. We are currently using RASMI to explore user responses to party political television broadcasts and envisage future use in the area of online market research.

Information visualization, temporal media, assessment, uncertainty

1. INTRODUCTION

Broadly, user interfaces designed to capture user emotions or user opinions can be classified into static interfaces, used to capture collections of data at a single point in time and dynamic interfaces used to capture continuous flows of data over a period of time. [1, 2].

In computer-based environments the predominant approach to has been to devise visual interfaces that are static in nature and allow the user to indicate choices by interacting with interface components such as check-boxes, radio buttons, text fields and pull down list boxes. Interface components are arranged to enforce a high degree of structure over the interaction process, requiring the user to make absolute decisions regarding choices, views and opinions. This is best illustrated by multiple choice-based interfaces which do not allow the user to express any degree of uncertainty regarding choices made or any degree of preference between offered choices. For example if a user is asked to make a choice from three different options A, B and C using a traditional multiple choice based interface, the user cannot indicate that they are split between choice A and C but prefer A over C. Furthermore a user cannot indicate that they are 70% sure option A is correct. More generally, incorporating degrees of uncertainty relative to other choices is not possible. When using static interfaces to allow users to assess or evaluate temporal media presentations users are generally restricted to providing input after the presentation has been shown rather that during its playback.

Dynamic interfaces, which may be software or hardware based, are used to capture flows of data input over time [3, 4]. These interfaces are often used when users are required to provide feedback on temporal media presentations, such as party political broadcasts, television advertisements or musical compositions. Historically, dynamic interfaces have two dimensions representing opposing options or choices and users are asked to oscillate using a slider control between the two dimensions indicating their view. Slider movements are recorded and then mapped on to the time-line of the media presentation. Following this mapping it is possible to deduce a user's view on a presentation at a particular point in time. Although such interfaces are effective in situations where only two choices exist, a significant limitation is their failure to address situations where a third view point exists or where there is a need to indicate choice relative to other possible choices. Furthermore since it is possible for the user to make many movements of the slider whilst viewing a media presentation, the amount of data generated by dynamic interfaces can be relatively large and often requires detailed analysis before it can be interpreted. The value of using information visualisation techniques to interpret large datasets is known and well understood [5, 6].

We present a novel equilateral triangle-based interface that dynamically captures user emotions or user opinions in real time. The interface enables users to express these emotions or opinions with degrees of uncertainty. By moving the mouse over the triangle, where vertices represent different emotions or opinions users can indicate, for example, how much they agree, disagree or do not know, when asked to express an opinion or register an emotion whilst watching a video presentation. We have placed special emphasise on the process of transforming user inputs into visual form thereby enabling the viewer to observe, browse, interpret and understand the information captured. 2D Interactive graphs are used as an engaging backtracking tool enabling easy identification of when within a media presentation and with what degree of uncertainty a user wished to register an emotion. Information captured is also visualized by means of an Equilateral Triangle Scatter Graph and line graphs.

2. BACKGROUND AND MOTIVATIONS

The real-time advanced spatial measurement interface (RASMI) outlined in this paper is in part based on the work of Moore [7]. In [7] Moore proposes a Spatial Probability Measure technique as a way of measuring response certitude. Response certitude is the level of confidence a person has in their opinion. The technique involves having a triangle with three views at each point. A question is asked of the user and the answers are displayed on each corner of the triangle. The user then moves the mouse to a point within the triangle and the mouse distance from the three points is recorded. The three distances from each point help to determine the correctness of the user's response. The closer the mouse is to a corner point the higher the user believes in a certain opinion. Moore's approach utilises a static interface in which users are presented with a question or statement and asked to indicate one of three answers on a static form. In contrast the work reported here aims to explore how to design a human computer interface that supports user input across a vector field of encompassing three sets of data points in real time. We wished to allow users to use a mouse to input a collection of data points whilst viewing a digital video file, of for example a party political broadcast, that represented their degree of emotional uncertainty relative to other available choices. For example, given three different emotional states A, B and C we wished to allow a user to express that they felt 70% like the emotional state A and 40% like emotional state B and 15% like emotional state C. Furthermore, given a choice between A, B and C. allowing the user to express that they are not going to choose C for certain but that their decision is divided between A and B, that they preferred A over B and I are 80% convinced by A.

We were further motivated by the need to develop an architecture for the interface that would allow a users input to act as parameters for locating points of interest with the digital files playback timeline. We wished to create an intuitive display in which a snapshot of users' emotions could easily be observed. The inability of many of the current generation of assessment and evaluation interfaces to enable users to express emotion incorporating degrees of uncertainty relative to other choices was a further motivating factor for our research.

3. REAL-TIME ADVANCED SPATIAL MEASUREMENT INTERFACE (RASMI)

The user interface of RASMI is shown in Figure 1. The interface consists of: A persistent media player that allows direct access to the timelines of digital video files played; A triangular interface capable of recording three distances from the point where mouse was clicked to each of the vertices, thus allowing judgment of conflicting emotions or opinions and a information visualisation interface used to graphically represent individual and group user input to help identify key emotional responses and important trends. User and group data can be saved in session files for future processing by a set of information visualisation graphics.

Via the information visualisation interface user input from recently performed sessions can be viewed using a set of 2D Interactive graphs that link user responses to points in media playback timelines. 2D Interactive graphs act as a backtracking tool enabling easy identification of when within a media presentation and with what degree of uncertainty a user wished to register an emotion. Information captured is also visualized by means of an Equilateral Triangle Scatter Graph and line graphs.

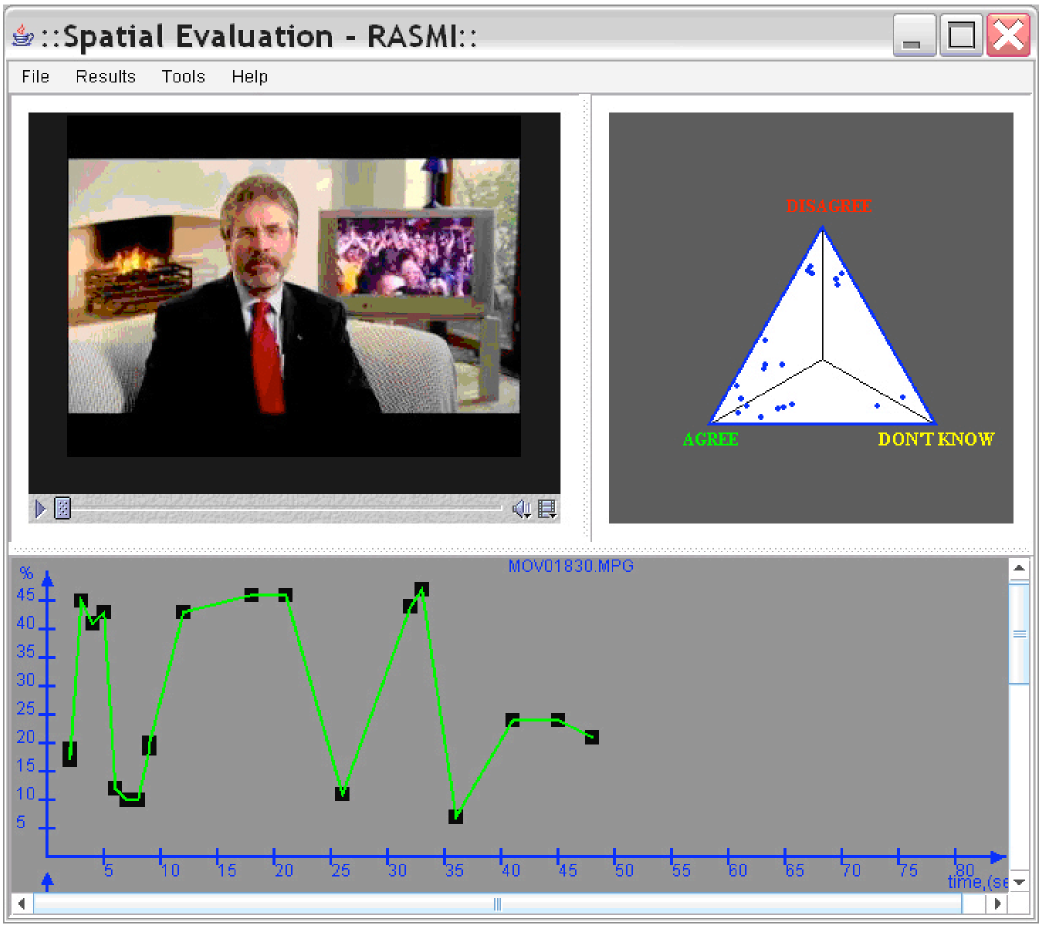


FIGURE 1: RASMI Data Visualisation after the end of evaluation session

3.1 Equilateral Triangle User Interface

User input is captured via the triangle input system shown in Figure 2. This is a graphical representation of an equilateral triangle which is used to designate an area within which the user can move the mouse. Each corner of the triangle has an opinion such as agree, disagree and don't know. The user clicks play on the media player or inside the triangle to start the video file. The user then moves the mouse near the corresponding corners of the triangle where a certain opinion lies. While the user is moving the mouse the location of the mouse pointer within the triangle is being recorded for every second of the video file. The location of the mouse pointer at any one time can then be measured against the distances from the three corners of the triangle each showing a conflicting opinion. This gives three distance values which can be used to calculate degrees of uncertainty when ever user input is received.

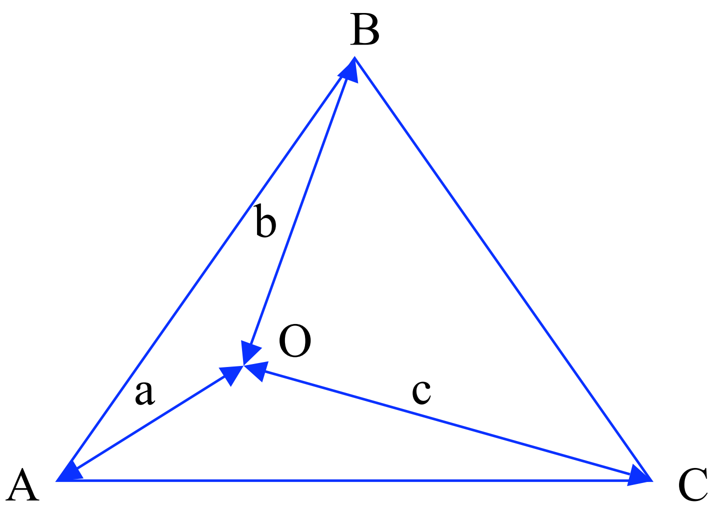


FIGURE 2: Equilateral Triangle Input System

In this Figure 2, segments a, b and c represents distances AO, BO, OC accordingly, from imaginative clicked point O to each of the vertices A, B and C. As sum of segments a, b and c with each click can vary, a percentage calculation is used to compute a level of uncertainty towards a particular vertex. The spatial measurement formula for each vertex is as follows:

$$A_u = \frac{a}{a+b+c} * 100$$

$$B_u = \frac{b}{a+b+c} * 100$$

$$C_u = \frac{c}{a+b+c} * 100$$

Where A_n, B_u and C_u are measurements of uncertainty corresponding to each of three vertices (%). It therefore follows that the level of uncertainty is low then the user is surer of his opinion, that if it was high.

3.2 Data Visualisation in RASMI

Data visualisation is the process of transforming numerical or symbolic data into a graphical representation that provides a more meaningful level at which users can interpret data. Within RASMI, data is visualised using 2D Interactive Graphs, Triangular Scatter Graphs and arithmetic mean Bar Charts. The 2D Interactive Graph shown in the bottom panel of Figure 1 is one of three generated during a RASMI session. Each graph plots the emotional responses associated with vertices. Where user has registered a change in emotional state, a media access point is generated and displayed. By clicking on this point a user can directly access and review the point in the temporal media presentation shown at the time of the change in emotional state. The Triangular Scatter Graph shown in Figure 1 provides a trace of all points clicked within the triangular interface during a session. This provides a snapshot of user behaviour.

4. CONCLUSIONS

The RASMI outlined in this paper is part of a larger research effort [8,9] into the development of real-time dynamic interfaces for interruption and evaluation. The purpose of RASMI is to demonstrate how to design a real-time interface that allowed users to engage with temporal media and express imprecise Intuitive responses. We hope that use of spatial measurement techniques presented here, provides a novel approach to allowing users to express emotion incorporating degrees of uncertainty relative to other choices. Future development of the system will involve providing more scope for users to trace different emotional states via a Hexagonal Input System.

REFERENCES

[1] E. H. Chi, 2000, A taxonomy of visualization techniques using the data state reference model, INFOVIS '00: Proceedings of the IEEE Symposium on Information Vizualization, IEEE, Computer Society. page 69.

[2] B. Shneiderman and C. Plaisant. 2004, Designing the user interface: strategies for effective human-computer interaction. Addison-Wesley Longman Publishing Co., Inc.
[3] C. Chen, editor, 2004. Information Visualization Beyond the Horizon. Springer ISBN 1852337893.
[4] M. Weber, M. Alexa, and W. Muller. 2001, Visualizing time-series on spirals. In INFOVIS '01: Proceedings of the IEEE Symposium on Information Visualization 2001 (INFOVIS'01), IEEE Computer Society page 7.
[5] D. A. Keim, 2001, Visual exploration of large data sets, Communications of ACM, 44(8): pages 38–44.
[6] E. R. Tufte. The visual display of quantitative information. Graphics Press, Cheshire, CT, USA, 2001.
[7] D. R. Moore, 2005, Exploring gender effects on the spatial probability measure. In Proceedings of the 5th IEEE International Conference on Advanced Learning Technologies, ICALT 2005, IEEE Computer Society, pages 435–436.
[8] James Ohene-Djan, Andrew Sammon and Rachel Shipsey, 2006 Colour Spectrum's of Opinion: An Information Visualisation Interface for representing degrees of Emotion in Real Time,10th International Conference, Information Visualisation 5, 6, 7 London , England (Accepted)
[9] James Ohene-Djan and Rachel Shipsey, 2006 , E- Subtitles: Emotional Subtitles as a Technology to assist the Deaf and Hearing-Impaired when Learning from Television and Film, the 6th IEEE International Conference on Advanced Learning Technologies ICALT, (Accepted)

Virtual Human Modelling and Animation Through a Sketching Interface

Chen Mao, Sheng Feng Qin, David K. Wright
School of Engineering and Design, Brunel University, Uxbridge, Middlesex, UB8 3PH
Chen.Mao@brunel.ac.uk, Sheng.Feng.Qin@brunel.ac.uk, David.Wright@brunel.ac.uk

Virtual beings are extensively utilised in today's public entertainment, while ordinary users are rarely involved in their modelling and animation process, due to the lack of appropriate expertise, equipment, and computer skills. We developed a new method and novel sketching interface, enabling everyone who can draw to interactively "sketch-out" 3D virtual humans and animation. We devised a "Stick Figure→Fleshing-out→Skin Mapping" graphical pipeline, which decomposes the complexity of figure drawing and considerably boosts the modelling and animation efficiency. We developed a gesture-based method for 3D pose reconstruction from 2D stick figure drawings. We investigated a "Creative Model-based Method", which emulates a human perception process to transfer users' 2D freehand sketches into 3D human bodies of variational body sizes, shapes and fat distributions. Our current system supports character animation in various forms, including articulated figure animation, 3D mesh model animation, 2D contour animation, and even 2D NPR animation with personalised drawing styles. Moreover, sketch-based crowd animation and 2D storyboarding of 3D multiple character intercommunication are also supported by this system. Our interface is developed through a user-centred design process. A preliminary user study (questionnaires and sketching observation) was conducted to support the overall system design. An informal user study was undertaken to evaluate the early stick figure animation interface. Our latest full figure sketching interface has been formally tested by various users through performance tests, sketching observations, and interviews.

Virtual human modelling and animation, sketching interface, storyboarding, user centred design, user study

1. INTRODUCTION

Virtual beings are extensively utilised in today's public entertainment (i.e. 3D games, Hollywood films, multimedia), while ordinary users are rarely involved in their modelling and animation process, due to the lack of appropriate expertise, equipment, and computer skills. We developed a new method and novel sketching interface, enabling everyone who can draw to "sketch-out" 3D virtual humans and animation. This paper presents the highlights of our approach and the interactive experience of using our sketching interface.

2. SKETCH-BASED HUMAN MODELLING AND ANIMATION PIPELINE

Natural figure sketches are featured by many 'noises' – foreshortening, contour over-tracing, body part overlapping, shading and shadow, etc. To decompose the complexity of direct 3D modelling and animation from fully rendered sketches, we designed a "Stick Figure [1]→Fleshing-out→Skin Mapping" pipeline (Illustrated in Figure 1). This is inspired by the drawing sequence recommended by many sketch books and tutorials. Meanwhile, in principle, it echoes the modelling and animation pipeline in commercial packages (i.e. 3DS Max, Maya). Regarding our current design, functionalities at different levels were gained for different users. Thus they can choose to make simple stick figures, create delicate 3D surface models, or explore further to animate these sketch-generated creatures. Moreover, models can be exported to commercial packages at any level to be refined by their function kits. Figure 6 shows our virtual human sketching interface and some produced 2D and 3D animations.

A preliminary user study (including *Questionnaire* and *Sketching Observation* study) was conducted to explore design questions, identify users' needs, and obtain a "true story" of figure sketching. 60 questionnaires were collected and processed; and 14 sketching observations were conducted with various users (including artists, designers, animators, graduate students, and researchers) in pre-defined drawing scenarios. The study outcomes were adopted to develop a natural and supportive sketching interface.

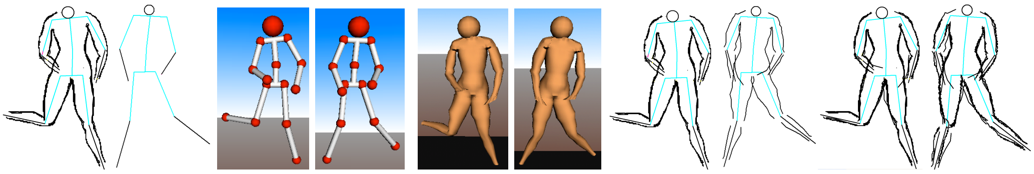


FIGURE 1: Users first draw stick figure key frames to define a specific motion. Then, they can "flesh-out" any existing stick figure with body profiles to portray an imaginative character. The system can automatically "perceive" the body size (skeleton proportion) and shape (body profile and fat distribution) from the sketched figure, and transfer it into a 3D virtual character. The resulting 3D skin surface can be mapped onto each of the posed stick figure key frames, which can be further interpolated as the final 2D and 3D character animation.

3. SKETCH-BASED 3D STICK FIGURE MODELLING AND ANIMATION

We developed a sketch-based gesture interface [1], which enables users to "draw" 3D stick figure animations. An on-line drawing assistance is provided by our interface for real-time figure proportion and foreshortening maintenance. We developed a "Multi-layered Back-front Ambiguity Clarifier", which utilises figure depth rendering gestures, human joint Range of Motion (ROM), and key frame coherence to identify user intended 3D poses from stick figure drawings. Since a quick and imprecise sketching may accidentally generate physically impossible poses, we offer a "Figure Pose Checking and Auto-correction" routine to detect ill-posed body parts, highlight them, and give proper corrections based on human body joint ROM and balance (See Figure 2(Left)). Our system supports an interactive design process, through which 3D figure models can be viewed and continuously updated in response to user's incremental sketching (See Figure 2(Right)). Once obtaining a series of 3D stick figure poses, user can easily sketch out motion paths and timing, and add their preferable music/background. The resulting 3D animation can be automatically synthesized in VRML after a single user click (See Figure 3). We conducted an informal user test on this stick figure animation interface to evaluate its functionalities and usability. The study results have been incorporated into the later full figure sketching interface design.

FIGURE 2: (Left) An original drawing with its ill-posed and auto-corrected 3D figure models; (Right) After incremental sketching, the figure pose is changed according to the modified perspective rendering.

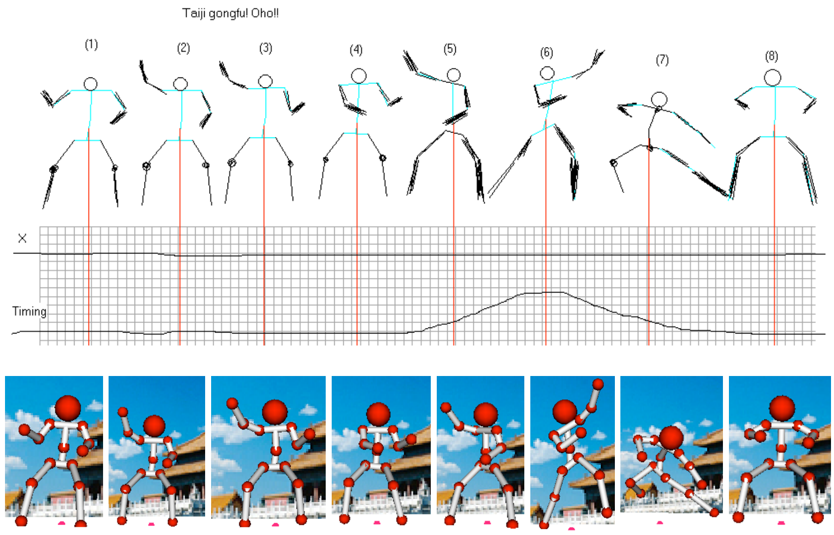


FIGURE 3: Animation for Taiji performance: The timing curve is given separately for controlling a long animation. A touch sensor (pink ball) is located inside the world for triggering the music.

4. FLESHING-OUT TO CREATE VARIATIONAL HUMAN BODY MODELS

Founded on the stick figure interface, we developed a full figure sketching interface [2], which enables users to sketch-out and animate virtual humans of variational body sizes, shapes, and fat distributions. As introduced earlier, users can depict the visual appearance of a virtual character through "fleshing-out" a single stick figure with body profiles. Improved on [3], which enables sketch-based spherical object modelling, we investigated a "Creative Model-based Method", which emulates a human perception process to model irregular/complicated human skin surfaces. It can perceive the body size (skeleton proportion) and shape (body fat distribution) of a sketched figure and transfer it into a plausible 3D counterpart model, through continuous graphical comparisons and generic model morphing (See Figure 4 (Left)). Our generic model (See Figure 4(Right)) is a three-layered (skeleton-fat tissue-skin) anatomical model (The Visible Human Project®), which can be transformed sequentially through rigid morphing, fatness morphing, and surface fitting to match the original 2D sketch (See Figure 5). Moreover, we offer an auto-beautification function to regularise 3D asymmetrical bodies from users' drawing imperfection (See Figure 5(c)). Users can interactively refine their 3D models by over-sketching 2D figure profiles (See Figure 5(d)). Modifications can be made on any key frame sketch, to achieve the updated 3D model. In addition, a post-processing function is provided for varying an existing figure model by changing its body proportion. The sketch-based modelling of human heads/hands/feet is not enabled at this stage, which is also a challenge for other related approaches.

5. SKETCH-BASED MULTI-LEVEL CHARACTER ANIMATION

Our interface [2] supports character animation in multiple levels, including articulated figure animation, 3D mesh model animation, 2D contour animation, and 2D personalized NPR animation (See Figure 1 and 6). Following the "Stick Figure→Fleshing-out→Skin Mapping" pipeline, a 3D virtual human animation is accomplished by wrapping the sketch-generated skin surface onto a series of posed stick figures, which can be further interpolated via VRML with the associated graphical motion definition. Figure 6(Top right) shows snapshots of the 3D animation specified by Figure 6 (Left), which provides an insight of our virtual human sketching interface and its intuitive graphical tool kits. 2D NPR animation is generated by successively interpolating the extracted key figure contours (with the mapped fleshing-out contours) and playing their in-betweens on sketching interface. This approach is different from the traditional Cel animation, and users do not need to render each key frame once a single key figure is fleshed-out. Figure 6 (Bottom right) shows snapshots of a 2D NPR animation in a doodled countryside view.

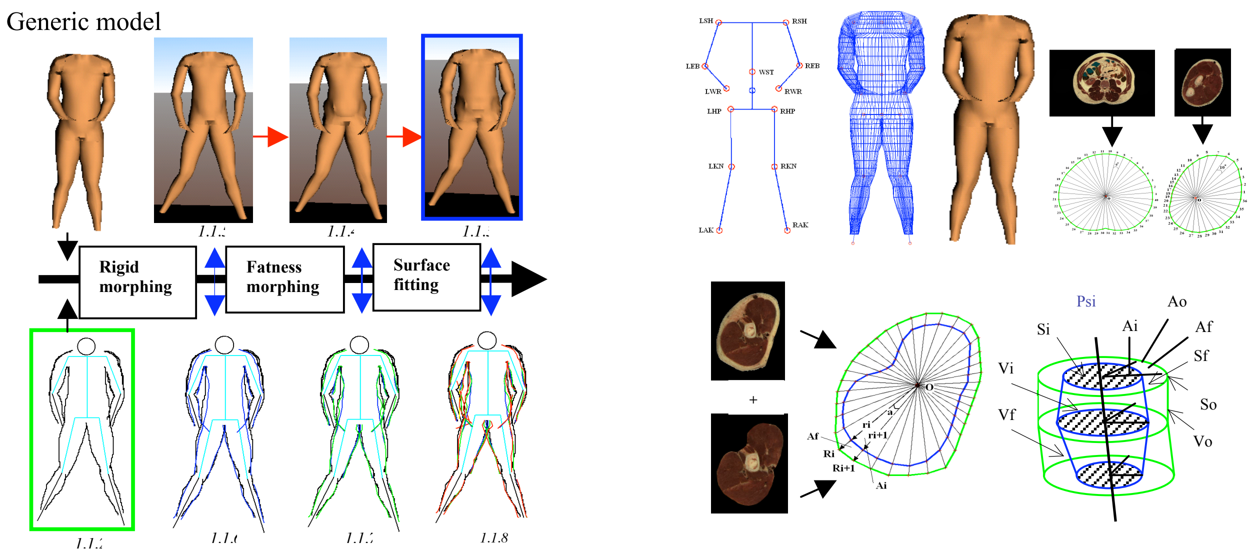


FIGURE 4: (Left) Transfer a 2D freehand sketch into a 3D body through an automatic system perception, graphical comparison, and generic model morphing process; (Top right) Template skeleton; wireframe generic model; rendered generic model; ray casting cadaveric cross-section images to digitise the 3D generic model; (Bottom right) Ray-casting each pair of cross section images (with and without fat layer) to digitise the template fat distribution.

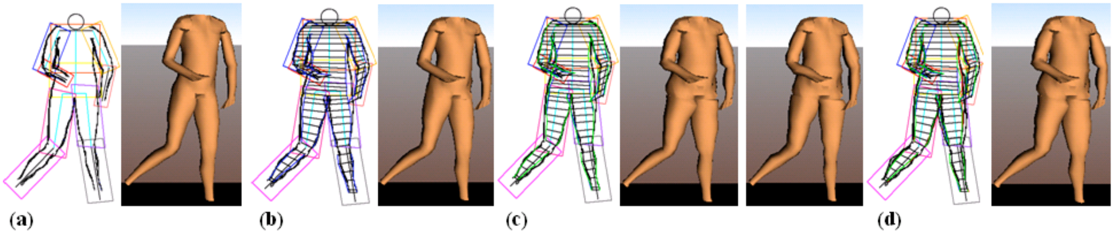


FIGURE 5: (a) The input 2D freehand sketch and the 3D model after rigid morphing; (b) Graphical comparison to get the fat distribution measurements and the fatness morphed model; (c) Graphical comparison to get the surface fitting measurements; the model with and without system auto-beautification; (d) Overtracing body contour (right lower torso, and lower legs) to modify an existing 3D surface model.

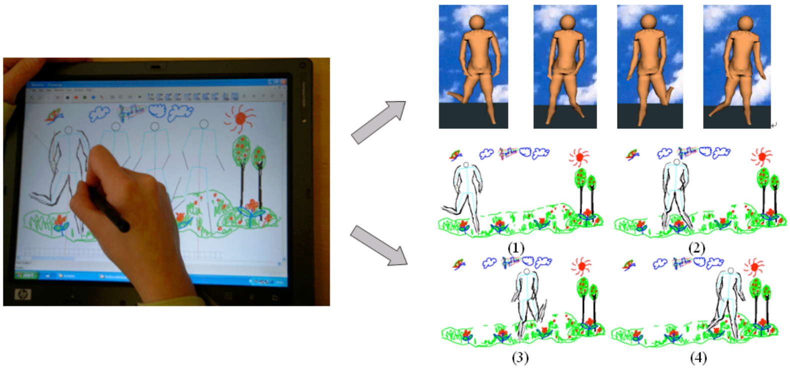


FIGURE 6: (Left) The user is sketching-out a virtual human and its motion on a Tablet PC; (Top right) A sketch-generated 3D dancing character; (Bottom right) A 2D NPR animation played on the sketching interface in a doodled countryside view.

6. SKETCHING-OUT CROWD ANIMATION AND CHARACTER INTERCOMMUNICATION

In our system [2], users can build their own 3D character and motion library, and animate a population of virtual humans through motion retargeting and a sketch-based actor allocation in 3D space (See Figure 7). Moreover, users are able to storyboard character intercommunication in each story scene (by either stick or full figure drawing). The system can deliver an immediate 3D scenario, in which virtual actors are acting and intercommunicating with each other (See Figure 8).

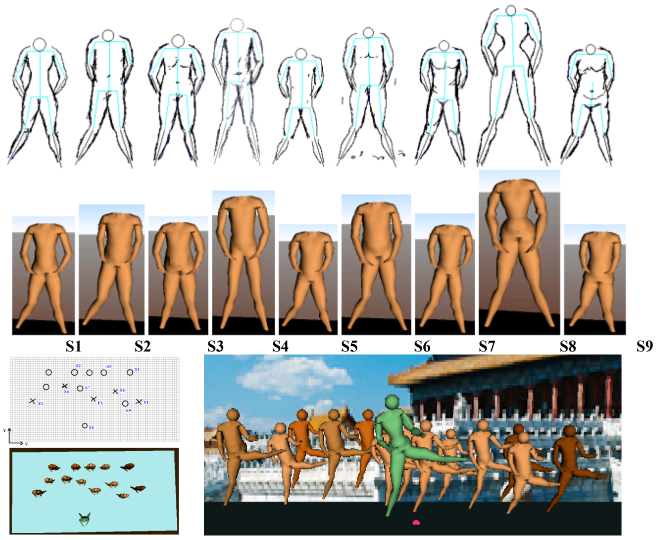


FIGURE 7: (Top) A variety of 3D virtual humans and the original drawings by different users: an artist (S3), a design student (S4), animators (S6, S7), graduate students (S1, S2, S5, S9), and a child (S8); (Bottom left) 2D top view plan with the denoted character locations and the corresponding 3D result; (Bottom right) Kungfu group animation with music and background.

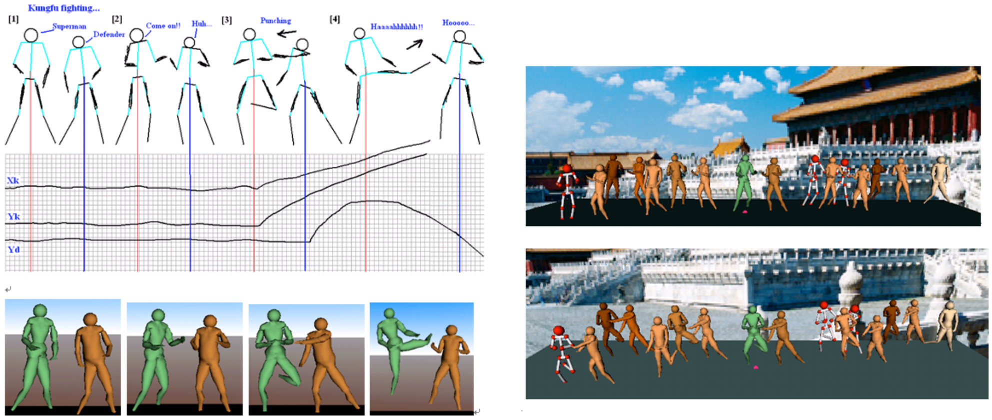


FIGURE 8: (Top left) The Kungfu fighting storyboards with associated dialogues and motion curves (The motion timing was act-out by the curve drawing speed.); (Bottom Left) Two sketch-generated 3D characters are fighting with each other; (Right) A crowd of sketch-generated virtual humans and stick figures are fighting with each other in 3D virtual world.

7. USER EXPERIENCES

On the completion of the current virtual human sketching interface, we evaluated it formally with various users (including an artist, an animator, design students, a 12-year-old boy, etc), through *performance tests*, *sketching observation*, and *user interviews*. After a short tutorial, users rapidly learned the modelling process, and began to sketch-out their own virtual human animation (on Tablet PC) within minutes (*6.27* and *6.75* minutes on average for a *3-frame stick* and *full figure animation* respectively). Through sketching observation and interviews, our sketching interface is proved by users to be easy to learn and use, and enjoyable for fast virtual human modelling and animation. Figure 7 (Top) shows the sketches and the variational human bodies created by users during the test, which have been integrated into two group animations, shown in Figure 7 (Bottom) and 8 (Right).

8. CONCLUSION

Human modelling and animation is a recognized challenge and labour-intensive task, which has been, until now, confined to the domain of professionals. This research draws on existing drawing skills of ordinary users to create and animate their own characters through 2D freehand sketching. In the future, we will adhere to our user-centred approaches and continuously improve our sketching interface according to users' needs.

REFERENCES

[1] Mao, C., Qin, S.F., and Wright, D.K. (2005) A sketch-based gesture interface for rough 3D stick figure animation. Proceedings of Eurographics Workshop on Sketch Based Interfaces and Modeling 06, 175-183. Dublin, Ireland.
[2] Mao, C., Qin, S.F., and Wright, D.K. (2006) Sketching-out virtual humans: from 2D storyboarding to immediate 3D character animation. To appear in proc. ACM SIGCHI International Conference on Advances in Computer Entertainment Technology 06, Hollywood, Los Angeles.
[3] Igarashi, T., Matsuoka, S., and Tanaka H. (1999) Teddy: a sketching interface for 3D freeform design. *Proc. SIGGRAPH '99*, 409-416.

HCI@FACT: Artists on Usability Exhibition

David England[1], Marta Rupérez[2], Caen Botto, Josh Nimoy, Simon Poulter
[1] Department of Computing and Mathematical Sciences,
Liverpool John Moores University, Byrom Street, Liverpool. L3 3AF
[2] FACT, 88 Wood St, Liverpool L1 4DQ
d.england@ljmu.ac.uk, ruperez@fact.co.uk, universomente@xicnet.com, josh@jtnimoy.net, simon@viral.info

The Human Computer Interaction @ FACT exhibition asks a group of three new media artists to engage with aspects of interaction. Their responses have been enjoyed by the public via a series of workshops and an exhibition at the FACT Centre, Liverpool. The workshops and exhibition aimed to provide members of the public with a different view of interaction than the standard, mouse, keyboard and screen.

The main exhibition consists of; Caen Botto's CD03: the mechanics of the spell, a hyper-musical instrument; Josh Nimoy's Mixed Hello, a set of camera vision interaction pieces; and Simon Poulter's AKM, Automated Knowledge Machine, a play on the ATM and internet knowledge access. Here we explain the pieces further and report on our current experiences with the exhibition. The interactive experience will recreate one or more of the installations. Some of the pieces and video clips of their use can be seen at http://www.hci-fun.org.uk.

Non-standard interaction, new media art, art works

1. INTRODUCTION

The Human Computer Interaction project is an EPSRC-funded public engagement with science initiative. Whereas most HCI research aims to inform designers and other researchers, this project aims to introduce HCI concepts to the general public. The subversive aim of the project is to create a guerrilla army of knowledgeable people who will demand more of their interactive experiences, and complain when they are presented with sub-standard interaction. The project is a collaboration between the first author at the School of Computing and Maths, Liverpool John Moores University and FACT, the Foundation for Art and Creative Technology. JMU and FACT began a search for three artists working in new media who were willing to engage with the concepts of HCI and produce interactive art pieces. Below we present the work of the three artists and our experiences through workshops and the exhibition.

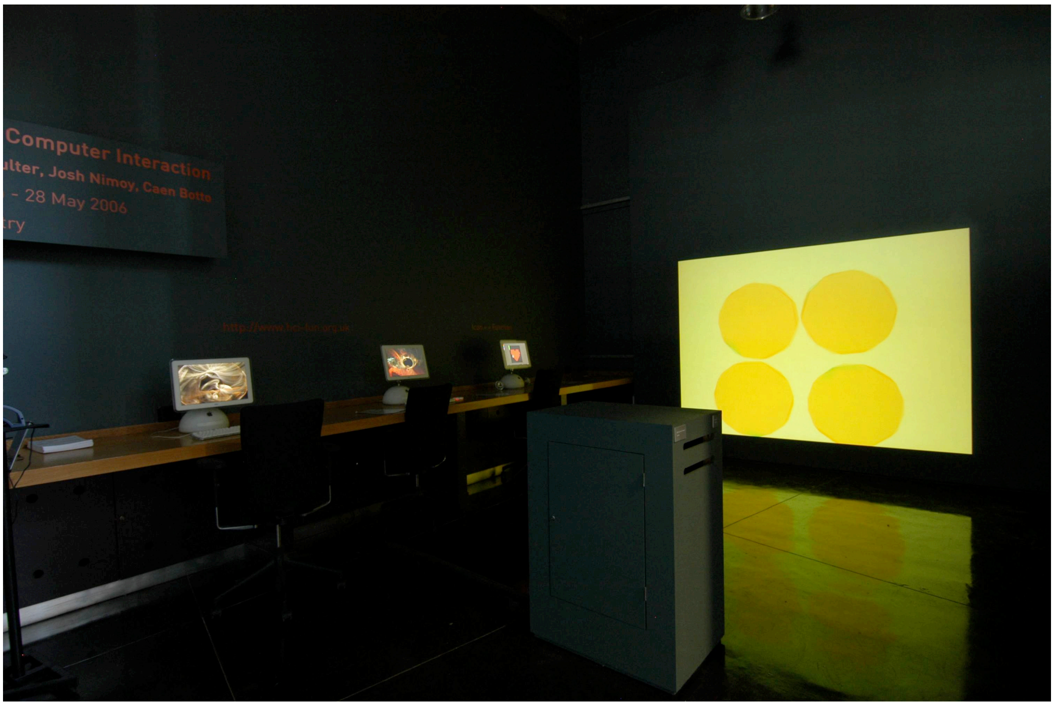


FIGURE 1: FACT Media Lounge and the HCI Exhibition

JOSH NIMOY

Josh Nimoy presents two pieces, *icon==function* and *Mixed Hello*.

Icon==function builds on Josh's previous work, the viral, waste of time, *Balldroppings* and is also a response to the following observation by Sine McDougal

> "Icon interpretation is inherently ambiguous because the relationship between icon and function is not determined by a set of well-defined syntactic and phonological rules.", *Coping with Complexity Workshop, Bath, 2004*

Icon == Function explores the intuitiveness of graphical user interfaces: how certain icons and symbols that we use on our computer screens have been historically and 'logically' assigned certain meanings. Users can use the six pieces of icon=function to create a visual layout the produces different soundscapes.

Mixed Hello is a Mixed Reality installation that gives the users the feeling of interacting with a projected screen using their shadow. The mechanism behind the interaction is based on Camera Vision[1] technology: the projected computer screen frame – a projected software environment and objects - is synchronized with the frames captured by the web-cam, matching the corners together. This way, any shadows or other activity detected by the web cam view can be matched to the virtual world as if each camera pixel were one mouse cursor, clicking the screen. The different pieces produced a vast range of interactions with children especially being willing to engage with the pieces. They were played with the different ideas of pace of interaction, 2 1/2 D interaction (making their shadows bigger and smaller), and moving between competitive and cooperative play as the different pieces supported single or joint use of the projected shadows.

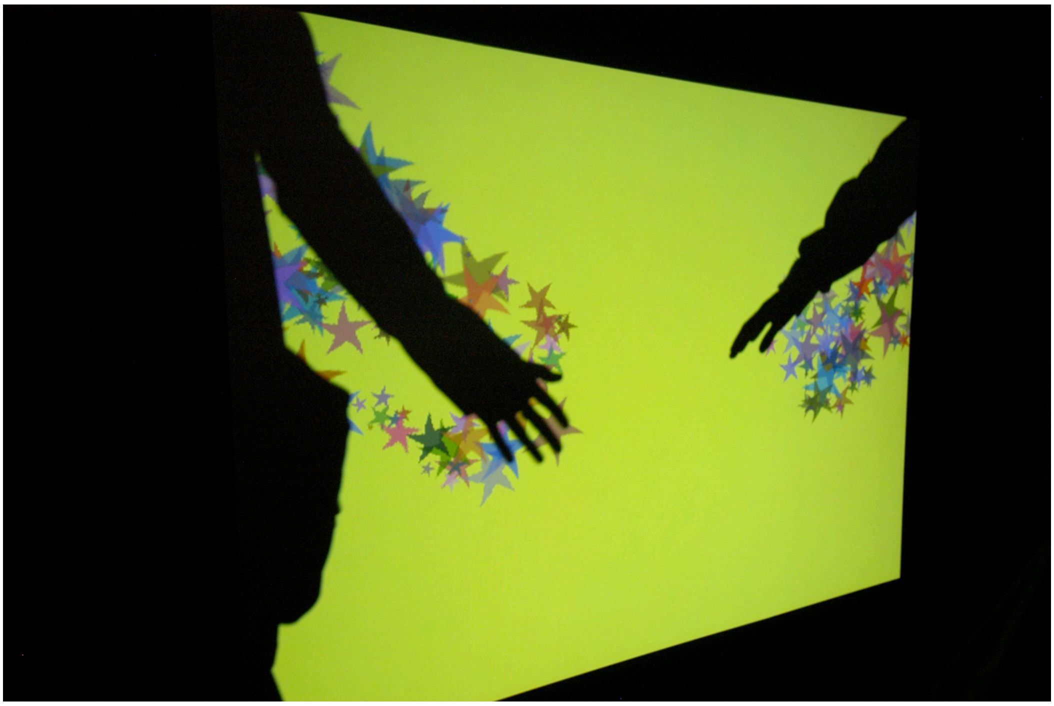


FIGURE 2: Mixed Hellos STARS

SIMON POULTER - AKM

The *Automatic Knowledge Machine* is Simon's latest humorous take on humans' relationship with technology. It presents the – seemingly - infinite pool of knowledge available online as the ultimate reason for the Internet.

All this infinite knowledge, available at the users' finger-tip, has been structured around certain key moments of the history of ideas: Global markets (currency and language systems), iconoclasts (from Jesus to Mark E Smith), Tesla (science, genius, electrical equipment and systems), history repeating itself, child's play, communications (from fire to mobile technologies), Liverpool, humans (how humans work psychologically) and music.

[1] A technique of Mixed Reality pioneered by Myron Kruger in the 70s

FIGURE 3: Automated Knowledge machine

On his writings about perspective, Leone Battista Alberti recommended that painters thought of the canvas as a window to the world. Simon's *AKM* reinforces the idea of the computer screen as "windows" open to all the knowledge in the universe…

Relying on the immediacy of operating a system through a touch-screen, the evolution of the *AKM* would recreate the very intuitive act of accessing knowledge in our brain. Ideally, in our cyborgian future, we would all wear implants in our brain directly connected to Google – to all the knowledge on the Web.

Cash machines have long been used in HCI research and teaching. For example they have been used to get over ideas of **task modeling** and **goal-oriented analysis**; people approach interactive situations with a plan in mind which can be broken down in sub-goals and actions. Here, the main goal is to get money. The sub-goals are to enter the card, enter the pin, enter the amount and retrieve the card. Early ATM's broke the user's cognitive model because they gave the user their money before returning the card - the users had achieved their main goal and discarded their sub-goal: people would walk away from these machines leaving their bank card in the slot. So, for the last couple of decades all machines have given back the card before the money to keep in step with the user's mental model of what should happen. Users do not seem to notice that Simon's AKM never returns their cards.

CAEN BOTTO _ CD03 MECHANICS OF THE SPELL

Caen's pieces involved sound and movement interactions which were fed into a Max/MSP program which produced organic visualisations and sounds. In the first workshop a series of floor pads were used to trigger sounds and visualisations projected on the surrounding walls. Again children led the way in innovative interactions by rolling over the pads, jumping on the in unison and performance a variety of dances on the connected pads. In the exhibition setting the sounds and visualisations were trigger by two sources of ambient sounds; first from the gallery space itself; secondly from web visitors whose microphone inputs also affected the visualisation. Thus Caen has created a hyper-instrument capable of being played by numerous, remotely located participants.

FIGURE 4: C0D3

Acknowledgments

This project was funded by the EPSRC Public Engagement with Science Initiative. We are grateful to Matt Phil, Dane Watkins, MetalCulture and all the staff at FACT for their support. Images, courtesy of Brian Slater.

REFERENCES AND WEB LINKS

[1] Sine McDougall and Martin Curry, “More Than Just a Picture: Icon interpretation in Context”, proceedings of *Coping with Complexity Workshop, Bath, 2004*
[2] Project Website and videos www.hci-fun.org.uk
[3] Josh Nimoy, www.jtnimoy.net
[4] Caen Botto, http://universomente.xicnet.com/
[5] Simon Poulter, http://www.viral.info

Posters

Investigating the Communication of Emotions Through Multimodal Technologies and Gestures

Ravi Kuber, Suziah Sulaiman and Ann Blandford
UCL Interaction Centre, University College London, UK
r.kuber@qub.ac.uk; {s.sulaiman, a.blandford} @ucl.ac.uk

In human–human interactions, touch is an important means of communicating emotional content. Touch-based technologies may have a role to play in augmenting computer-mediated communications, such as Chat and Instant Messaging. The study reported here investigated whether emotions could be effectively understood via haptic and gestural interfaces. Each interface was developed to communicate a specific emotion, based on established descriptions of the visual and gestural representations of positive and negative emotions. Sixteen subjects were asked to identify the emotion being expressed by reference to 'emo-cards' which displayed a larger range of emotions. Participants successfully identified the emotion of surprise, were partially successful in identifying happiness and disgust, but were unable to reliably identify sadness. We outline a programme of further research to determine which emotions can be reliably communicated, and how to design effective representations for communicating those emotions.

Haptics, gestural interaction, emotion, multimodal interaction

1. BACKGROUND

Users of real-time communications technologies such as Instant Messaging and Chat are increasingly communicating informal and emotional message content [7]. However, due to the lack of sensory cues surrounding transfer of data, additional time is spent explaining and clarifying emotions, in a way that is not needed in face-to-face communication. One method of improving the current state of exchange of emotional information could be through the sense of touch, which could help overcome the loss of subtle non-verbal communication cues, such as pauses in speech and the use of gesture. Previous studies have examined the benefits of using touch as a method for augmenting the communication experience of using instant messaging [5, 7]. Each of these studies was based on the premise that by using properties of the haptic modality, a richer and more stimulating interaction can be gained, as touch can provide key information about user state and the situational and social context. Haptic feedback is considered to provide an effective mechanism for online collaboration, and is thought to be helpful in asynchronous communication [7]. Gesture is also believed to play an integral role in this process. It was hypothesised that study participants would be able to perceive and recognise the emotions 'designed into' a set of haptic cues that were combined with gestures, and that these effects would represent emotions in a more meaningful and representative manner than current methods, such as verbalising feelings and emoticons allow.

2. SELECTION & DESIGN OF GESTURES & HAPTIC CUES

Four interactions were designed to represent different positive and negative emotions (surprise, happiness, disgust and sadness). Each required the user to perform specified gestures, while receiving particular haptic feedback from the system via the PHANToM Desktop device stylus. Emotions were selected from a set of six culturally independent basic emotions identified by Ekman [2]. The remaining two emotions from that set (anger and fear) were omitted in order to approximately balance the numbers of positive and negative emotions under investigation. The designs of the gestures were based on those devised by Paiva et al [6] for an affective control toy called SenToy, used to control a synthetic character in a computer game; the resulting gestures and haptic cues are summarised in Table 1. The design of haptic feedback was based on a study by Moody et al [4], who found that smooth material communicated positive feelings of happiness and sleekness while rougher textures conveyed feelings of uneasiness, discontent and confusion.

3. EXPERIMENTAL PROCEDURE

Sixteen participants with no prior experience of force-feedback were provided with introductory training on the PHANToM device, and were presented with a series of visual and haptic interfaces, to solicit feelings regarding textural and other haptic effects. Each participant then worked with each of the four emotional

Basic Emotion	Haptic cues & gestures devised for the study
Surprise	A large level of force-feedback (vibration) is produced when user initially moves the Phantom stylus towards an object located on the interface. Adapted from [6]
Sadness	The object on the interface can be slowly moved away from the user using the stylus. The haptically-modelled effect of friction has been used, to produce an effect of resistance. Adapted from [6]
Happiness	Using the stylus to move the object upwards, using small and rhythmic movements. Motion should not be constrained. An object with a light perceived mass can be moved in an environment free of gravity. Adapted from [6]
Disgust	Moving an object away from the user with a slight increase in tension (spring stiffness), whilst squeezing the stylus button. Adapted from [6]

TABLE 4: Emotions & gestures

interfaces, presented in a randomised order to minimise the occurrence of learning effects. Each user was asked to perform the gesture specified in Table 1 with the corresponding interface, and to verbalise any emotion they felt as a result of the interaction. Following each interaction, the participants were invited to specify the emotion by either selecting one from the emo-card based on that of Desmet et al [1] using basic emotions from Ekman [2], or by naming one of their own choice. They were also encouraged to discuss the strength of the emotion felt. If no emotion could be verbalised, participants were asked to discuss the intensity of the force-feedback, and whether the haptic interactions led to a positive, negative or neutral state. A semi-structured questionnaire covering topics such as the role of haptic cues and gestures, issues of engagement and modality integration was also presented.

4. RESULTS & DISCUSSION

The main hypothesis aimed to address whether emotions could be conveyed using haptic cues and gestures. Results from the study generally tended to support it. When questioned, only two out of the sixteen participants believed that emotions could not be communicated using haptic cues and gestures. Findings indicate that some emotions are easier to communicate multimodally than others. Pre-tests had revealed that the haptically-rendered surface texture of smoothness evoked feelings of calmness, whereas the texture of roughness created feelings of slight discomfort and uneasiness. When identifying individual emotions, the interface designed to communicate the emotion of surprise was accurately named by fifteen out of the sixteen participants, confirmed by binomial testing revealing a high level of significance ($p<0.0005$). Haptic cues and gestures conveying emotions of happiness and disgust were also identified by the majority of participants. However, binomial distribution testing of interfaces for happiness ($p=0.196$) and disgust ($p=0.122$) revealed that these responses could have been determined due to chance. Further trials would be need to be conducted to isolate whether the haptic and gestural feedback were solely responsible for producing these emotions. The interface designed to communicate sadness was found to be ineffective, potentially attributed to sadness being a longer-lasting emotion [2]. Participants spent a matter of seconds interacting with the interface, rather than taking their time to experience the emotion. As emotions do not occupy distinct categories or have structured boundaries [2], it is very common for negative emotions such as anger, fear, disgust and contempt to be labelled as being in the same group. The results suggest that a negative emotion was perceived by users, but they were unsure of which one. The secondary hypothesis aimed to examine whether a natural and engaging environment be produced using these cues. Results revealed that the process of interaction with multimodal interfaces, created a rich and stimulating experience for participants. Haptic cues were found to be natural and balanced in combination with visual feedback experienced.

5. CONCLUSION & FURTHER WORK

This study has provided evidence that haptic technologies and gestures could be used to convey emotion, providing informative, meaningful cues. Our findings suggest that surprise was significantly better communicated than other emotions. This study has delved into the issues of communicating individual emotions, which have not been addressed in previous studies of this nature. Findings have provided a basis for further study. As a next step in the research project, it will be necessary to carefully rework the haptic cues that were not successful in communicating the intended emotions through a series of iterations. For example, a lower level of resistance provided by pushing an object, may have helped users to identify the emotion of sadness. Following the redevelopment of cues, the focus will shift to real-world applications such as chat. A future study should monitor how users interact with haptics and gestures over a longer period of time, examining any shared conventions that develop, and whether these conventions lead to users expressing their emotions more effectively than using conventional text-based chat. The work presented here indicates that such investigations are likely to be fruitful.

ACKNOWLEDGEMENTS

We thank Paul Cairns, Pieter Desmet, Agnès Guerraz and Peter Greasley for their input to this work.

REFERENCES.

[1] Desmet, P.M.A., Overbeeke, C.J. and Tax, S.J.E.T. (2002) Designing products with added emotional value; development and application of an approach for research through design. The Design Journal, 4, 32-47.
[2] Ekman, P. (1999) Emotions revealed. Times Books, New York, NY.
[3] Maclean, K. (2000) Application-Centered Haptic Interface Design. In Human and Machine Haptics. MIT Press, Cambridge, MA.
[4] Moody, W. et al (2001) Factors underlying fabric perception. Proceedings of Eurohaptics 01, Birmingham, UK.
[5] Morris DiMicco, J., Lakshmipathy, V. and Fiore, A.T. (2002) Conductive Chat: Instant Messaging with a Skin Conductivity Channel. Proceedings of CSCW 02, New Orleans, LA.
[6] Paiva, A. et al. (2002) SenToy in FantasyA: designing an affective sympathetic interface to a computer game. Personal and Ubiquitous Computing, 6, 378-389.
[7] Rovers, A.F. and van Essen, H.A. (2004) HIM: A framework for haptic instant messaging. Proceedings of CHI 04, pp.1313-1316.

An Empirical Study of a Question-Based Authentication Technique

Ann Nosseir, Richard Connor and Crawford Revie
Department of Computer and Information Sciences, University of Strathclyde
Livingstone Tower, 26 Richmond Street, Glasgow, United Kingdom G1 1XH
FirstName.SecondName@cis.strath.ac.uk

Stolen or forgotten passwords and ID cards can limit users' accessibility to and interaction with their personal devices or email and consequently limit users' mobility. Further, biometric assessment and ID cards are costly in terms of equipment, installation and maintenance. Decisions on selecting authentication techniques tend to involve a trade-off between security and two other dimensions: mobility and cost. In low risk situations, where the implications of unauthorized access are limited, using an expensive technique cannot be justified. The use of passwords is thus the most common solution as they are the cheapest technique; however they have memorability problems. In this paper, we investigate a novel question-based authentication technique that uses electronic personal histories to generate questions. The aim is to investigate the possibility of differentiating between genuine users and impostors based on their respective answers to these generated questions. We have also investigated the impact of psychological parameters on memorability.

Security, identity, human memory, passwords, usability

1. INTRODUCTION

There are situations that require a reasonably secure authentication mechanism with minimum personal effort. For example, users would like to access their emails from any computer either in a public internet café or from their home. They need an authentication technique that it is simple and mobile; they need a mechanism that they don't have to carry around and will not be forgotten, and which they can use irrespective of their location. Other examples could include accessing certain physical domains such as students' labs, common rooms or sports halls. These situations require a technique that does not impose unnecessary cost on the organization. The existing techniques such as "who you are" (biometrics) or "something you carry" (passports) are expensive to deploy. "Something you know", such as passwords, are the cheapest option but have memorability problems [2]. In the area of security usability, researchers have suggested alternative mechanisms which are either recall-based using facts and opinions, or recognition-based using images [1,5,7]. In an earlier study, we investigated a novel idea using electronic personal history which is well known only to the individual and whose volume is too large for impostors to learn [3]. In the earlier study, we reported poor results based on memorizing random questions extracted from personal electronic calendars. Thus, there was a need to better understand how the human memory worked. In psychology human memory may be classified as follows: sensor memory, working memory and long-term memory. Long-term memory is divided into episodic, procedural and semantic memory. In our research, we have focused on the long-term memory and in particular episodic memory which some researchers define as "autobiographical memory", and which postulates that human beings remember recent, repetitive and pleasant events better than other types of event [6].

2. EXPERIMENT

In this experiment, we used personal web sites because they are publicly available and at this stage we are simply exploring potential hypotheses. We have used the personal websites of twenty-four staff members, chosen from a set of colleagues well known to each other in a UK university department, to generate eight different questions. Four pairs of questions were generated for each of the psychological types [6] (recent, repetitive, location, and pleasant) with each pair consisting of a Boolean (true / false) and a four-option multiple choice question. Participants were divided into three groups of eight members. In addition to answering their own authentication questions, each participant was asked to act as an 'impostor' by attempting to answer all 56 questions (8*7) for the other seven members in his/her group.

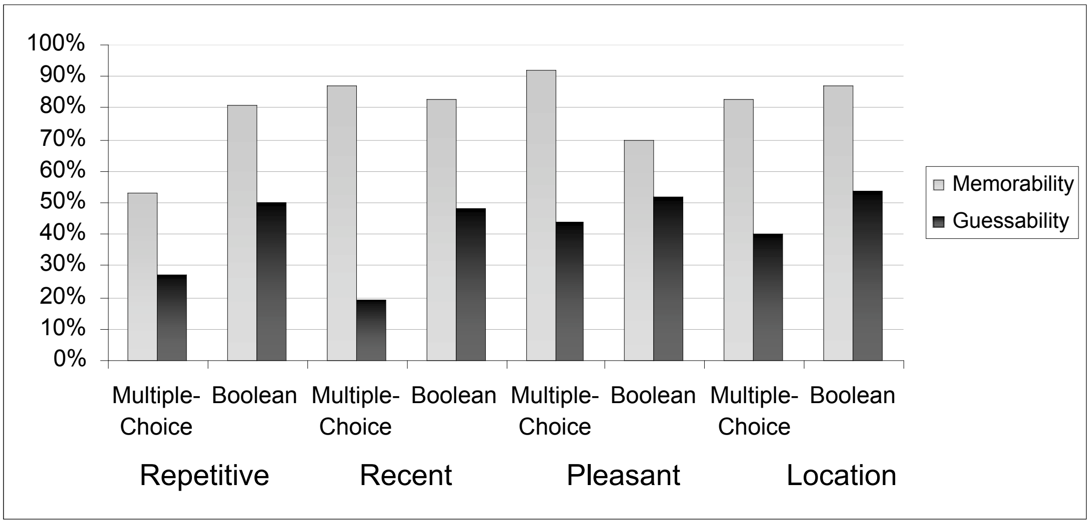


Figure 1: Histogram of Comparing Memorability and Guessability for Each Question Type

3. RESULTS

Statistical tests indicate that there is a significant difference ($p < 0.01$) between the participants' and impostors' answers. Overall the percentage of correct answers provided by the genuine participants was 80%, compared to 43% achieved for impostors. However, perhaps of more interest were the differences between the participants' and impostors' answers over the various types of questions, as shown in Figure 1. Of the eight question types, the genuine participants were best able to answer those classified as "pleasant" with multiple choice options (92%), while in the case of questions being answered by impostors the lowest score was obtained for "recent" with the multiple-choice option (19%). This is lower than the 27% reported by Zviran [7] for cognitive password techniques which use open questions about personal facts and opinions. In addition the "recent" question type with multiple choice led to a high score on average for genuine participants (86%) and consequently the largest difference between the two types of user. In general you might expect the impostors to score around 50% and 25% for the Boolean and multiple choice questions respectively (based on chance and the fact that there were four options). While this appears to be the case for impostors on the Boolean questions, it only holds true for the "repetitive" and "recent" question types using multiple choice.

4. CONCLUSIONS

The results from this novel approach provide additional elements to the authors' earlier findings [3], and illustrate the benefits of using a variety of question types. The approach adopted within this research can be extended to other sources of electronic data, such as that stored on mobile phones or sensor data. Further investigation is required to gain confidence in the results and to explore additional parameters that significantly distinguish those seeking genuine access from impostors.

REFERENCES

[1] Brostoff, S. and Sasse, M.A. (2000) Are Passfaces More Usable Than Passwords? A Field Trial Investigation, HCI2000 conference, Proceedings of HCI2000 Springer. Sunderland, U.K., September, pp. 405-424,

[2] Jianxin, Y., Blackwell, A., Anderson, R. and Grant, A. (2004) Password Memorability And Security Empirical Results in Cambridge, IEEE Security & Privacy, September, pp. 26-31.

[3] Nosseir, A., Connor, R. and Dunlop, M. (2005) Internet Authentication Based on Personal History – a Feasibility Test. Proceedings of ACM world wide conference. Tokyo, May.

[4] Stajano, F. (2002) *Security for Ubiquitous Computing*. WILEY: University of Cambridge, UK.

[5] Takada T. and Koike H. (2003) Awase-E: Image-based Authentication for Mobile Phones Using User's Favourite Images, Conference HCI03, Udine, Italy, Proceedings of Mobile HCI03 Springer, 8-11th September, 347--351. http://www.w3.org/TR/WCAG10/

[6] Wagenaar W. (1986) My Memory: A study of autobiographical memory over six years, *Cognitive Psychology, 18,* 225-252.

[7] Zviran, M., and Haga, W. (1990) Cognitive Passwords: The Key to Easy Access Control, Computer Security, 9, pp.723-736.

An Experimental Interactive Application Managing Cultural Data, Based on Customizable User Interface Design

George Pehlivanides
Department of Cultural Technology and Communication
University of the Aegean, Tirtaiou 1, 81100 Mytilini, Lesvos, Greece
g.pehlivanides@ct.aegean.gr

This poster presents an interactive game through which users could test their knowledge about the cultural identity of Lesvos (an island in the northeast of Greece). The application was implemented during the academic year 2005 by the students of the Department of Cultural Technology and Communication of the University of the Aegean, in collaboration with the academic staff of the module "Cultural Representation I". The objective of the research procedures which took place during the implementation phases was the exploration of cultural representation practices through the use of new technologies and in particular through Customizable User Interface design. To emphasize the physicality regarding the modes of interaction with the application, a custom built floor map with embedded sensors was especially developed. On this floor map, users were able to interact with the content by walking on geographically arranged areas, and simultaneously by answering questions related to various aspects of the cultural identity of the island.

Cultural representation practices, customizable user interfaces, cultural information interaction design

1. INTRODUCTION

Using alternative ways of interacting with computer systems that outreach common computing practices - such as the use of a keyboard / mouse / monitor - has nowadays become quite ordinary in various technological sectors. Especially, in cases that involve the analysis / negotiation of cultural content - the data management of which is very often restricted by the use of conventional interaction methods - experimental approaches based on physical computing could open the way for the establishment of new forms of communication practices. From a technological point of view, the easiness of interconnecting natural and virtual spaces using special hardware and relative interaction design developing toolkits [1, 2], offers the possibility to experiment with a variety of expression modes. Having this supplementary ability, the design team is able to design and develop interaction environments integrating elements of familiar communication practices met in everyday life. Within an interdisciplinary framework, consisting of methodological approaches from the knowledge fields of Cultural Sciences, Applied Informatics and Information Interaction Design, this poster presents the development phases of an interactive application managing cultural data, based on Customizable User Interface design.

2. DESCRIPTION OF THE APPLICATION

The multimedia application entitled "Walking on Lesvos" consisted of two components: a) an interactive floor map of 40 sq. m. developed and used as an abstract representation of Lesvos, and b) a Graphical User Interface (GUI), which maintained the functionality of the game by projecting in large scale format all the relative game components (such as images, text, video extracts, assistant symbols, etc). The users, separated in teams, starting from the enter point of the floor map, had to follow routes indicated by special interactive key points which represented cities, villages, as well as areas or monuments of specific historical socio-cultural and environmental interest. The aim of the game was to reach the final destination of the floor map without failing to answer a series of questions related to each geographically specified area. Users had to follow specific routes, or alternatively, according to the rules of the game, they could choose to follow different pathways, in order to gain additional points. The questions were presented in the form of multiple choices; in case of selecting the wrong answer the team would lose a round.

3. INFORMATION INTERACTION DESIGN

While investigating users' interaction modes with the presented cultural content, emphasis was given to sensorial design issues [3], and particularly to physicality: Apart from the sensorial stimuli of audiovisual character provided by the multimedia components of the application, also tangibility was examined by exploring sensorial approaches through touch, kinaesthetics and the sense of direction / location in space. More specifically, users could interact with the thematic units of the application based on two sequential actions: a) by walking to - stepping on the key points of the map, and b) by choosing the correct answer with the help of a handheld wireless device. Additionally, a series of custom-built highlight indicators with changing brightness and rate corresponded to each key point of the user interface supported the unfolding of the plot. According to the answers of each team, the patterns of the routes were formulated by visual indications, contributing in this way to the cognitive process. Furthermore, the surrounding space was changing colour (green, red or blue) depending on the team, thus denoting its identity within this space. In combination with the above mentioned design practice, the development of the application in the form of a game, offered the possibility to use elements of familiar communication practices such as competition, participation, engagement, emotional involvement, etc. which constitute an ideal methodological tool for knowledge acquisition.

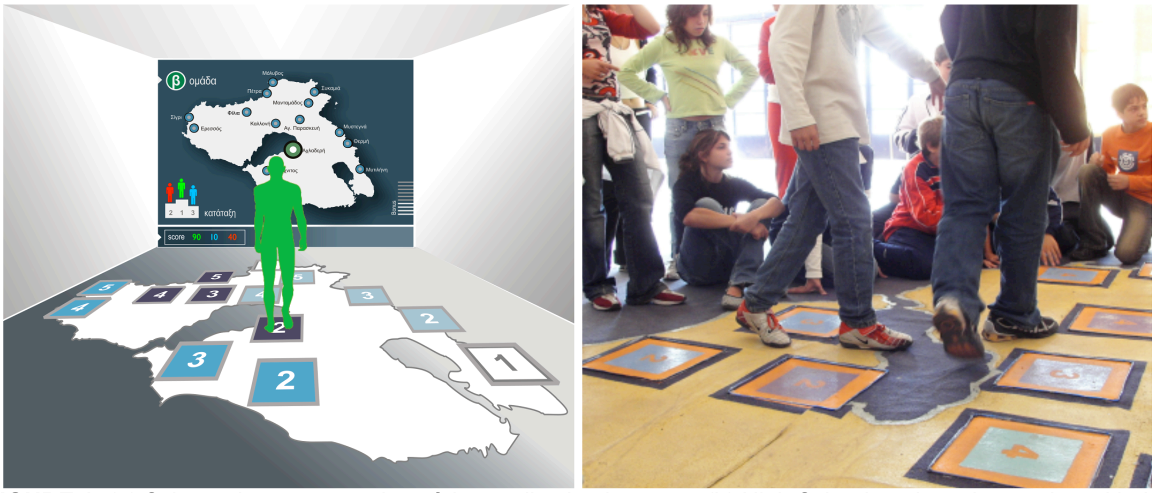


FIGURE 1: (a) Schematic representation of the application in space, (b) High School students interacting with the application

4. CONCLUSIONS AND ISSUES FOR FURTHER INVESTIGATION

A pilot version of the application was presented in June 2005 within the framework of a multimedia exhibition and received positive feedback by the users, especially concerning its overall presentation. In September of the same year it was presented to a targeted audience composed of high school students participating at an educational programme in a high school of Mytilene (the capital city of Lesvos). This gave the opportunity to examine the effectiveness of the application as a learning tool for cultural heritage in its geographical context. Based on recent learning theories, it is currently accepted that physical means of interaction with the content combined with a game strategy enhance learning, and therefore this might be applied to geo-cultural knowledge acquisition.

During both events, structured questionnaires were distributed to the participators in order to conduct a summative evaluation. The evaluation was aimed at measuring the application's impact on the users, and extracting conclusions for the improvement of prospective versions. Consequently, a questionnaire was elaborated according to the following concerns: a) general impressions, b) usability, c) overall presentation, and d) interactivity aspects. At this moment the questionnaires are at the stage of indexing - evaluation process. The first evaluation results concerning the use of new technologies for the presentation of cultural knowledge proved to be positive and thus encouraged us to proceed to next research stage which will analyse the efficiency of the platform compared to traditional ways of presentation and conventional software based tools.

REFERENCES

[1] Greenberg, S. and Fitchett, C. (2001) Phidgets: Easy development of physical interfaces through physical widgets. Proceedings of UIST 01, Orlando, Florida, 11-14 November, pp.~209-218, ACM Press.
[2] Yim, J-D, and Nam, T-J (2004) Developing tangible interaction and augmented reality in Director. Proceedings of CHI 04, Vienna, Austria, 24-29 April, pp.~ 1541-1541, ACM Press.
[3] Shedroff, N. (1999) Information interaction design: A Unified field theory of design, in Jacobson, R. (ed.), Information Design, MIT Press, Cambridge Massachusetts.

LiLiPUT: Lightweight Lab Equipment for User Testing in Telecommunications

Peter Fröhlich, Peter Reichl, Raimund Schatz, Lynne Baillie, Wolfgang Weinberger and Florian Hammer
Telecommunications Research Center Vienna (ftw.), Donau-City-Straße 1, 1220 Wien, Austria
{froehlich, reichl, schatz, baillie, weinberger, hammer}@ftw.at

User testing in mobile HCI is still mostly conducted in stationary labs, even though the importance of field data is widely acknowledged. Starting from related work and a set of six key requirements for more successful and widespread field-based user testing, we present our Lightweight Lab Equipment for User Testing in Telecommunications environment ("LiLiPUT").

LiLiPUT, mobile usability, mobile HCI, methodology, laboratory equipment

1. INTRODUCTION

Recent research has clearly demonstrated that standard stationary lab testing is not sufficient for mobile usability evaluation and HCI research. First, users engaging in mobile situations tend to have focuses other than the device in front of them, a situation quite different compared to an indoor laboratory setting [1]. Thus, in order to obtain ecologically valid results, usability tests for mobile devices should be conducted in typical usage situations outside the lab. Second, whenever mobile context factors matter for design or research questions (e.g. amount of interruptions or the input modality depending on the situation), field studies need to be conducted [2]. Third, some mobile interaction types such as spatially aware services simply cannot be evaluated satisfactorily indoors.

Different kinds of field research methods have been proposed in order to better understand and better match the nature of mobile contexts: ethnographic field studies, diaries and experience sampling [3]. However, such studies "in the wild" are still quite rare; the vast majority of mobile usability tests and also many HCI research experiments for mobile applications are still carried out in stationary labs. One important reason for this is the additional effort required for test execution in an outside environment and the complexity of (manually prepared) data analysis required for most field research methods. Furthermore, the amount and reliability of collected data (e.g. video, audio, and observer annotations) tends to be lower than in stationary user studies.

2. REQUIREMENTS FOR FIELD-BASED USER TESTING EQUIPMENT

We have defined the following lab equipment requirements to enable more successful and widespread field-based user testing of telecommunications systems:

1. **Portability** (adequate size, power consumption for mobile equipment, etc.)
2. **Data richness and accuracy**: The collected data should be of a comparable level (e.g. comprehensiveness and accuracy) to that of a stationary lab (recorded handheld screen capture, video and audio data to document the test situation)
3. **Workflow efficiency**: The required workload for the research or consultancy team should not be higher than that of a traditional lab study. The data organisation and analysis should not consume more labour resources than when using a traditional lab.
4. Mobile **contextual data capture** (weather conditions, surrounding physical environment, gestures and emotional expression, etc.)
5. **Natural and seamless interaction**: Users should be able to move normally e.g. for example make telephone calls whilst engaging with the mobile device without any hindrances. To this end, they also should not feel uneasy or inhibited due to cables, obtrusive cameras or the surrounding team of observers
6. **Coverage** of all relevant mobile telecom application areas (not only mobile HCI, but also human-human communication and related QoS (Quality of Service) issues)

3. RELATED WORK

Mobile HCI research has only recently tried to meet some of the requirements outlined above. The most advanced solution in this regard was proposed by Roto et al [4]. In essence, 1 microphone signal and 4 camera signals are mixed and recorded on a DV recorder located in the user's backpack. The cameras capture the mobile device display, the user's face, the user's front (from the user's shoulder), as well as the user from the perspective of the observer (from a wireless camera on the observers shoulder. The proposed solution fulfils requirements 1-3 to a high degree, enabling a reasonable amount of mobility, data richness and accuracy, as well as a very high workflow efficiency (there is only one observer and the data is ready to be coded after the test). Furthermore, the video of the user's front partly helps to meet requirement 4, i.e. capturing mobile context. However, the obtrusive cables and cameras, mounted on the mobile device, may considerably hinder naturalness and seamlessness (req. 5). Also, coverage of all relevant telecom application areas, namely making telephone calls (req. 6) would be impossible.

4. THE LILIPUT APPROACH

The system presented by Roto et al [4] inspired us to develop the environment "LiLiPUT" (Lightweight Lab Equipment for User Testing in Telecommunications). In particular we were interested in meeting more closely the requirements of naturalness and seamlessness, whilst ensuring high standards in terms of data quality and accuracy. Therefore, we decided to develop a strictly wireless system, which relieves the user from carrying the cables and a backpack (see figure 1 below).

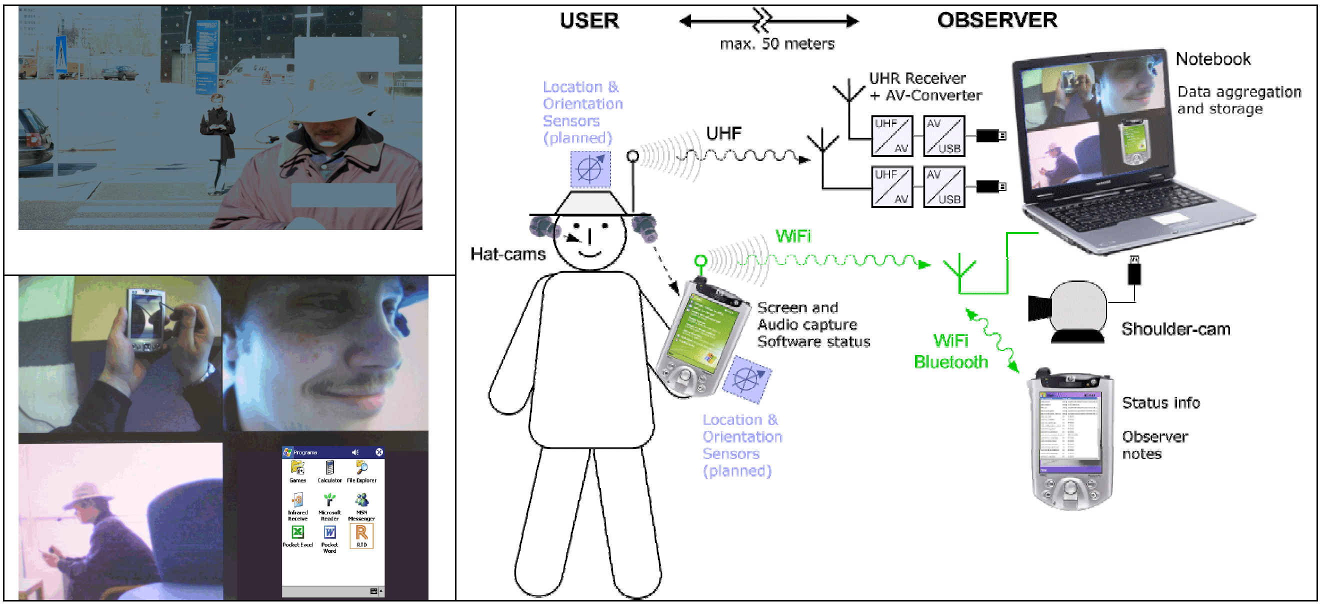


FIGURE 1: LiLiPUT System in the field (top left), captured video signals (bottom left) and system architecture (right). Features planned for the future are coded in blue.

The only part of LiLiPUT that the user is supposed to carry (apart from the mobile device of course) is a customised hat, on which two cameras, a microphone, wireless senders and batteries are mounted, which capture the user's line of sight and their face respectively. The face camera and its holder are removable for studies with strong requirements for seamlessness but without a special focus on facial expressions. The handheld screen display is captured via a remote display control software. The two wireless camera signals and the captured handheld display, as well as a web camera on the observer's shoulder are mixed and recorded on a notebook with software developed in-house running on it. In addition, the remote control software can be used to control the handheld display as required for undertaking Wizard-of-Oz studies.

5. ONGOING WORK AND OUTLOOK

As consequence of our first functional tests, we acquired a more powerful, waterproof notebook in order to achieve better recording performance and more versatile usage under different weather conditions. Initial user feedback that was informally collected during our functional tests, indicate that LiLiPUT seems natural and seamless to use, thus meeting req. 5. In the near future, we plan a systematic validation with regard to the requirements stated in this paper. We also aim to improve the automatic capture of additional mobile context data, such as time-stamped head and handheld orientation information gathered via embedded orientation and acceleration sensors.

6. ACKNOWLEDGEMENTS

This work is part of the project SUPRA funded by the Austrian competence centre programme Kplus. We would like to thank Antitza Dantcheva for her support with building up the hardware equipment.

REFERENCES

[1] Baillie, L. & Schatz, R. (2005) Exploring Multimodality in the Laboratory and the Field
Seventh International Conference on Multimodal Interfaces, Trento, Italy, September, 2005, Published by the ACM.

[2] Brewster, S. A. (2002) Overcoming the Lack of Screen Space on Mobile Computers. (2002) Personal and Ubiquitous Computing, 6(3), 188-205.

[3] Consolvo, S. (2003) Using the experience sampling method to evaluate ubicomp applications. IEEE Pervasive Computing, Special Issue on Human Experience, April-June 2003, 24-31.

[4] Roto, V., Oulasvirta, A., Haikarainen, T., Kuorelahti, J., Lehmuskallio, H. and Nyyssonen, T. (2004) Examining mobile phone use in the wild with quasiexperimentation. HIIT Technical Report 2004-1, 2004. www.hiit.fi/publicationsMobile eyetracking.

Usability Evaluation – Support for the Inclusion of Indirect Social Interactions

Robert Macredie and Emma Pickering*
School of Information Systems, Computing and Math's, Brunel University, Uxbridge. UB8 3PH, UK.
*{Robert.Macredie, Emma.Pickering}@brunel.ac.uk (*corresponding author)*

This paper argues that indirect social communication (i.e., the projection of a fashionable image) is a noteworthy part of use of personal and mobile technological artefacts and, as such, should be included in any evaluation of use. To explore this position, data was collected from within a larger ethnographically-inspired study which focused on exploring interactions of 'beyond the desktop' technologies in the real world. Four semi-structured qualitative interviews from this larger study were selected as having content relevant to indirect social communication and were analysed by interpretive phenomenological analysis. Of particular interest was the concern of fashionability, and the participants exhibition of tactics to manage communication. Conclusions drawn support the notion that the use of technological artefacts by a user is affected by indirect social interaction and so dimensions for indirect social interaction should be included in evaluations of use.

Usability evaluation, social interaction, personal computing, ubiquitous computing, interpretative phenomenological analysis

1. INTRODUCTION

There is a general awareness that technological artefacts, particularly those seen to fall within personal and mobile computing, are being used either intentionally or not, for indirect social communication – that is, social communication beyond the direct functional uses for which the artefact was originally designed. Examples of this indirect social communication include: use of the artefact as a prop to start a new scenario of social interaction, to interrupt or end an existing social interaction; or for social 'impression management' (see [2] for definition) through factors such as fashionability, or technological mastery and skill. Factors such as these appear to affect use in impacting upon when, where, and how, a user uses the technological artefact when in a social situation.

This paper thus posits that we must recognise that indirect social interaction is a noteworthy part of use, and so should be included as a dimension during evaluation of use.

2. TECHNOLOGY AND PRESENTATION OF SELF

When a user is using a personal and mobile technological artefact whilst in a social situation, that user is communicating messages about themselves to other people in the social vicinity (here, the term 'use' is used loosely to mean either active or passive use). Examples of these messages are introduced below:

- Technological savvy-ness and connectedness to the digital world may perhaps be projected through ownership and presence of an artefact on a person – having the 'right' gadget to assist in projecting a certain image;
- Knowledge, skill and mastery may perhaps be projected through an ability (or inability) of being able to use the artefact or its functionality;
- Personal taste and style may be communicated through aesthetics, such as style, type and looks of the artefact.

Impression management in social situations is well known [2]. Impression management through technological artefacts is less well known. Strom [5] identified many aspects of use of artefacts as social props, including the many indirect social messages that can be communicated. Traditional usability (as in the well accepted ISO 9241-11 [1] 'Guidance on usability' standard) however, mainly concerns performance and satisfaction measures. Although the standard does allow for identification of any goals pertinent to the usage situation, they are by nature thus intentional and known, and examples given focus on functional-type tasks (e.g., "…print copies of a report…" [1:4]). There is no specific mention or consideration within the standard towards any such indirect or unintentional social communication, particularly those that would generally not be considered ‘tasks’, such as projection of fashionability by a user leaving their phone out to be seen on their coffee shop table, for example. Thus the richness of this whole facet of use is being missed. In considering the integral nature of this facet to the use situation, ISO 9241-11 [1] in particular is thus limited in its application to use situations for personal and mobile artefacts used within social situations and this work seeks to begin to address this.

During data analysis from the aforementioned larger research project, many of these social messages are emerging. This project work thus seeks to follow up on this emergence and also expand the work by [5] to further identify and understand indirect social communication through these types of artefacts. The aims are to (i) identify instances of these indirect social communication messages, and (ii) identify categories of these messages, in order to then (iii) propose these categories for inclusion in evaluations of use.

3. METHODOLOGY

Work within this poster is extracted from relevant findings within a larger user-centred study that focused on exploring interactions of ‘beyond the desktop’ technologies in the real-world environment. Of the 9 semi-structured interviews available, 5 were rejected due to no mention of indirect social communication. The remaining 4 interviews were analysed firstly within case and then cross-case by interpretive phenomenological analysis (see [4] and [6] for technique), which is used to “…explore in detail how participants are making sense of their personal and social world.” [4].

3. FINDINGS AND IMPLICATIONS

Analysis is ongoing, however preliminary work has identified several interesting findings concerning users’ process or tactics in addressing this indirect social communication, and this is illustrated in two examples below:

An emerging theme throughout analysis was ‘fashionability’. One participant stated that he did not want an unfashionable phone because his friends would laugh at him. Of interest here is the *avoidance* of unfashionability. This appears to be more important to him than *encouraging* fashionability within the social context of his friends. Not only is fashionability a dimension of concern, but the nature of the negative form of the judgement will have implications in inclusion of dimensions – not just satisfaction but dissatisfaction, and the participant’s use of a ‘tactic’ - avoidance.

Context also had an impact, and one participant gave a particularly interesting example contrasting his handling of his inability to use a technological artefact in a work and in a home situation. At work the participant described how he would take time to form an ‘act’ in order to attempt to mask this inability, but summon assistance without his having to admit this inability. When at home, the participant stated he would just ask whoever was around. Thus the nature of the people in the vicinity affected how the participant dealt with the inability to use the artefact: he identified a different tactic for each context.

4. SUMMARY AND ONGOING WORK

Empirical work from this ongoing study has indicated the existence of indirect social communication in use of personal and mobile technological artefacts. It has identified that users themselves are engaging in impression management and also how their awareness of these issues impacts upon use. The empirical work therefore supports the argument that dimensions for indirect social communication are a consideration of use and so should be represented in evaluations of use. The main suggestion for further work would be for a study that specifically focused on investigating this phenomenon. As users demonstrated using tactics, the tactics scale presented by Lee et al. [3] may prove illuminating. This would assist in moving towards a more refined identification of these dimensions for consideration in evaluation.

It is unclear at this early stage how dimensions such as these will impact upon evaluation methodologies, particularly in considering the practicalities of how they may be included. It is suggested that further work, following the refined identification of dimensions, will then be required to establish *how* these issues may then be included in evaluations of use.

REFERENCES

[1] BSi (1998). BS EN ISO 9241-11:1998 Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11: Guidance on usability. CEN, European Committee for Standardization: 1-19.
[2] Gardner, W. L., Martinko, M. J., and Van Eck Peluchette, J. (1996). A self-presentational perspective of computer-mediated communications. In Beard, J. W. (ed.), Impression Management and Information Technology. Quorum Books, London.

[3] Lee, S-J., and Quigley, B. M. (1999). Development of a self-presentation tactics scale. Personality and Individual Differences 26:701-722.
[4] Smith, J. A., and Osborn, M. (2003). Interpretative phenomenological analysis. In Smith, J. A. (ed.), Qualitative Psychology: A practical guide to research methods. Sage Publications, London.
[5] Strom, G. (2002). Mobile devices as props in daily role playing. Personal and Ubiquitous Computing 6: 307-310.
[6] Willig, C. (2001). Interpretative phenomenology. In Willig, C. (ed.), Introducing qualitative research in psychology: advances in theory and method. Open University Press, Buckingham.

A Model for Structuring UML Class Diagrams to Support Non-Visual Interpretation and Navigation

Oussama Metatla, Nick Bryan-Kinns, Tony Stockman
Interaction Media & Communication Group
Department of Computer Science, Queen Mary, University of London. E1 4NS, UK
{oussama, nickbk, tonys}@dcs.qmul.ac.uk

Accessing textual information non-visually can be performed directly using Braille translators or text-to-speech engines. Problems arise when faced with graphically represented information, such as a diagram. For blind and visually impaired people, screen-readers are the most commonly used tools for accessing on-screen information, but they typically fail to deliver graphical information in appropriate and flexible ways. As a result, diagrammatic representations are inaccessible for such a population of users. We present the design of an experiment exploring a task-oriented approach for integrating speech and non-speech sounds to support the interpretation and navigation of conceptually relevant information encoded in a diagram.

UML diagrams, auditory representation, accessibility

1. INTRODUCTION

The inherent graphical nature of diagrams makes them partially or totally unusable by visually impaired and blind individuals. In areas such as Software Engineering and Systems Analysis & Design, the use of diagrammatic notation is fundamental to advancement, and the ability to interpret and generate such diagrams is now a trait of full professionalism. Within the growing community of object-oriented blind programmers, many feel that they miss out on participation opportunities in such disciplines as a direct result of accessibility issues related to diagrammatic notation[1].

We present the design of an experiment investigating the use of speech and non-speech sounds as an auditory representation for communicating conceptually relevant information that is encoded in a diagram. Our target diagrams are Class Diagrams; a widely used graphical notation for system modelling prior to commencing coding in the Unified Modelling Language (UML) standard.

2. BACKGROUND

2.1 Diagram Semantics

Cognitive science theories define diagrams as external representations that aid memory in problem solving tasks [3]. Some forms of diagrams represent entities that, even though graphically situated, successfully convey meaning regardless of their spatial and graphical properties. The graphical layout in such instances is simply an artefact of the presentation medium rather than an inherent part of the information [2].

It is claimed that by moving from sentential to diagrammatic representations, one can exploit additional dimensions for expressively representing information [3]. Such dimensions refer to the graphical properties of the syntax used in a particular diagrammatic notation. We believe it is possible to create similar dimensions by exploiting characteristics of the auditory medium of presentation in order to deliver comprehensible semantics of a diagram.

2.2 Non-Visual Diagrammatic Representation

2.2.1 Specialised Interfaces

A number of researchers have suggested alternative representations for accessing graphical information. A combination of tactile, synthetic speech and non-speech sounds were used in the Audiograf system [5] for instance. More recently, and with relation to Systems Engineering, the TeDUB system [4] used a specialised navigation interface to allow blind users to access semantic information of technical drawings. Evaluations of

[1] Information source: informal communication with blind programmers.

the system show that users successfully completed diagram-reading tasks, which included diagram exploration and information searching.

2.2.2 Navigational Strategies

Bennett [1] investigated navigational strategies specific to box-and-arrows diagrams. His findings show a correlation between the type of tasks to be performed and the navigation model employed to assist users perform. Similar ideas are investigated in the proposed work. We are interested in finding out the different interaction patterns and user behaviour when solving particular tasks in relation to a given mode of presentation.

3. EXPERIMENT

3.1 Design Approach

Our approach attempts to make up for the loss of dimensionality of expression when a diagram is described textually (or through spoken output) by extending it with non-speech auditory display. The experiment is thus designed to find out where such extensions are most effective. We are using a tree structure as an underlying model of structuring UML class diagrams. This model assumes that the most conceptually relevant components in relation to comprehension tasks are the objects, or classes, and their connections, or relationships. A given diagram is thus split into two main 'information-containers' an Objects component and Relations component. Figure 1 depicts this concept. The user can navigate the structure using the keyboard in a similar fashion to typical file menu browsers.

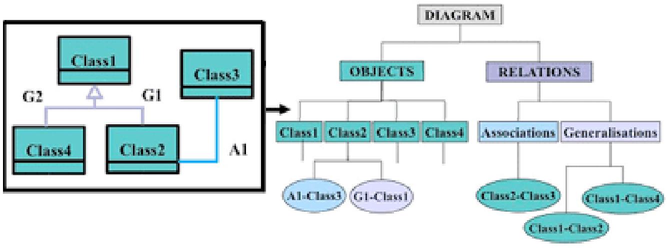


FIGURE 1: The class diagram (left) and the underlying structuring model (right). A tree structure is used to split a given class diagram into two main components.

3.2 Sound Design

Speech and non-speech sounds will be combined and played in parallel under user control to communicate information encoded in a given diagram through the structuring model. Presentation strategies will be implemented, the auditory display of which will be varied forming different experimental conditions to communicate either navigational information, contents of the tree structure, or both.

3.3 Methodology

3.3.1. Participants

a total of 15 subjects will participate in this experiment, each experimental condition will be tested by 5 subjects. A training session will be conducted for each participant in order to familiarise them with the particular display condition they will be using to solve a task.

3.3.2. Tasks

We will assess the efficiency of the proposed model in delivering a comprehensible structure of the diagram by testing it in relation to a set of user tasks. Its effects on the ease, as well as the accuracy, of user performance will be analysed. We are assuming that these tasks are necessary to gain a comprehensible interpretation of a given class diagram:

- Gaining an overview of the diagram, including the number of classes and relations it contains
- Examining the connections of a given class
- Examining the classes connected by a given relation

3.3.3. Observational method

User interaction will be automatically logged for later analysis and extraction of usage patterns and behaviour. At the end of the experiment, each participant will be asked to answer a questionnaire to gather qualitative data of their overall experience and subjective preferences. Each session will also be filmed for more in-depth analysis of the interaction.

4. CONCLUSION

There is a need for an understanding of how graphical representations can be translated from the visual to another modality while maintaining the same informational and computational affordances. Even though research into auditory display provides evidence of the potential of using sound in interface design, there is a lack of systematic studies for determining effective methods for creating auditory diagrams. Our work attempts to add on to the existing efforts to further our knowledge of how to support non-visual interaction with diagrammatic representations.

REFERENCES

[1] Bennett. J. D, (2002), Effects of navigation and Position on Task When Presenting Diagrams to Blind People Using Sound. Lecture Notes in AI, 2317:161.

[2] Brown. A, et al. (2004), Issues in the Non-Visual Presentation of Graph Based Diagrams. Proc. of the 8th Inter. Con. on Information Visualisation IV'04

[3] Cheng. P, et al. (2005) Cognitive Science Approaches to Understanding Diagrammatic Representations. Artificial Intelligence Review. Kluwer Academic Publishers, 2001.

[4] Horstmann. M, Hagen. Et al. CH. (2004), TeDUB: Automatic Interpretation and Presentation of Technical Drawing for Blind People. In CVHI2004 Conference & Workshop on Assistive Technologies for Vision and Hearing Impairment. Glasgow.

[5] Kennel. R. A., (1996), Audiograf: a Diagram-Reader for the Blind. Proc. of Second Annual ACM Conference on Assistive Technologies ASSETS'96, Canada.

Doctorial Consortium

Rethinking HCI for Information Fusion and Decision Support

Maria Nilsson
School of Humanities and Informatics
University of Skövde, Box 408, 541 28 Skövde, Sweden
maria.nilsson@his.se

The focus of information fusion has traditionally been on the technical aspects e.g. algorithms. However, the importance of also understanding human information processing is gaining increasing recognition. The research outlined here aims to incorporate and widen the perspective of HCI in information fusion and decision support research i.e. the utilisation of the user as an active component in the information fusion process. A new perspective of the information fusion process is proposed and distributed cognition is identified as a theoretical framework capturing the process. Also, initial results indicate that trust, uncertainty and temporal aspects are issues affecting the interaction between the user and the information fusion system.

Information fusion, HCI, cognitive science, distributed cognition, human factors

1. INTRODUCTION

Information fusion research originated within the military domain and refers, originally, to the process of combining information from multiple sources to perform inference which might not be possible with only a single independent source [1]. In other words, more precise knowledge can typically be gained by combining information from several sources. Due to the current development of complex information technology, the concept of information fusion is becoming increasingly important. This has been acknowledged and one could see that information fusion systems have had a growing interest and they can, now, be found in a wide array of domains typically characterised by large amounts of data and the need to take decision under time pressure. The University of Skövde has recognised this new exciting area and established a collaborative research program supported by the Swedish Knowledge Foundation together with 15 industrial partners. The program consists of seven different scenarios which cover areas such as retail, military and precision agriculture. The diverse usage of information fusion systems puts a new light on the involvement of the user i.e. the human decision maker. Due to the traditional technical focus of information fusion research, there is a lack of research covering cognitive constrains and human factors [1].

The research outlined in the following sections seeks to contribute by exploring the interaction between the information fusion system (i.e. decision support system utilizing information fusion) and the user.

2. PROBLEM STATEMENT

A common denominator between all information fusion systems is “the purpose” of the system i.e. the purpose to support the user(s)/the decision maker(s). It has been suggested that the overall efficiency of an information fusion system is strongly dependent on the user [1]. Therefore, there is a need to explore the interactions between the user and the information fusion system and to understand the factors which affect that interaction. Hence, the focus of this research is the development of a theoretical framework capturing the nature of information fusion processes and their embedding in the support of organisational, individual and automatic decision processes. In particular, the research covers organisational and cognitive demands and constraints that different types of information fusion systems have to face. In the end, the aim with the identified framework is to assist the development of future information fusion system and decision support system.

2.1 Contributions to HCI

An effective and efficient interaction between the information fusion system and the user is crucial [1]. However, today, there is a lack of research in the information fusion community focusing on HCI issues. The contributions of the research outlined will not only be to present, distribute and encourage the broad usage of HCI to a new audience, the information fusion community (i.e. the usage of HCI extends), but also explore the ideas of the user as an active component of the actual system. Thus, the results from this research might give new insight to be used and further developed by the HCI community. In particular, the research presented might change the view of the user’s contribution to the success of a system and enhance the incorporation of organisational factors. Also, the nature of the situations studied demands fast decisions based on displayed information representing complex information from multiple sources, today, this is central issues in HCI which can not be ignored. In the end, this project will be an example of how users can be utilised to achieve an overall effective information fusion process. As Hall and McMullen [1] point out: "*By rethinking the HCI for data fusion, we may be able to re-engage the human in the data fusion process and leverage our evolutionary heritage*" (p. 336). In my opinion, this could be an aim for the HCI community in general.

3. APPROACH

In order to develop a theoretical framework capturing the information fusion process one needs to perform a thorough examination of the subject i.e. the users' role in the information fusion process. The initial approach is to use all seven application scenarios involved in the information fusion research as case studies to get an overall picture of different users and different types of interactions. With an existing taxonomy developed by Blasch and Plano [2] as a starting point, the users in the different applications domains are analysed through interviews and case studies to understand the different requirements information fusion and decision support may have. This will result in the identification of different types of information fusion processes, different types of users and interactions, cognitive and organisational demands, and common terminology applicable across different information fusion domains. Also, we have proposed that *distributed cognition* [3] can be used as a methodology to capture the involvement of organisational, individual, and automatic decision processes to the actual fusion process and the interaction with the system, see [4]. Actually, the usage of distributed cognition in this research will further develop and enhance the proposition originally made by [5] who proposed that distributed cognition could be used as a new foundation for HCI.

In summary, guidance from an already existing taxonomy [2] is utilised, and, the knowledge gained is used with the help of distributed cognition to create a new framework capturing the actual *information fusion process*.

4. CONCLUSIONS

A literature study has resulted in the identification of known issues (e.g. trust, situation awareness, workload etc.) which influence the development and utilisation of an information fusion system (i.e. a decision support system utilising the power of information fusion). Also, an initial study has been carried out in order to capture the commonalities between the different projects within the information fusion research program. The initial conclusions from the study are that *trust*, *uncertainty* and *temporal aspects (of the decisions)* are issues applicable across most of the projects within the information fusion research program. Hence, these are issues which are likely to play an important role for the information fusion process and will be further explored in future studies. Also, the different projects have been classified according to the taxonomy presented in [2].

Furthermore, a first position paper [4] has been accepted for presentation at the Fusion 2006 -9th international conference on Information Fusion with the aim to not only present a *new perspective on the information fusion process* (i.e. a process which not only consider the actual system but also include the user) but also introduce *distributed cognition* as a methodology for information fusion. Also, a paper [6] regarding trust and situation awareness was accepted and presented at CHI´2005. The relevance of trust for information fusion systems will be further explored in future studies.

4.1 Future work

The fairly new area of Information fusion provides many interesting possibilities for a HCI researcher. The next step, for the research presented, is to further explore the different projects within the information fusion research program in order to create a framework capturing the information fusion process i.e. the role of the user. The plan for the upcoming year is to write-up the results from the studies performed for presentation at the fusion 2007 conference and for the decision support conference 2007.

ACKNOWLEDGEMENT

This work was supported by the Information Fusion Research Program (University of Skövde, Sweden) in partnership with the Swedish Knowledge Foundation under grant 2003/0104 (URL: http://www.infofusion.se).

REFERENCES

[1] Hall, D. and McMullen, S (2004). Mathematical Techniques in Multisensor Data Fusion. Artech House. London.

[2] Blasch, E. and Plano S. (2003). User refinements to aid the Fusion Process. Proc. of Multisensor, Multisource,
Information Fusion: Architectures, Algorithms, and applications, SPIE vol 5099.

[3] Hutchins, E. (1995). Cognition in the wild. MIT Press, Cambridge.
[4] Nilsson, M. and Ziemke, T. (2006). (*In press*) Rethinking level 5: Distributed Cognition and Information Fusion.
Proceedings of FUSION 2006, 10-13 July, Florence, Italy
[5] Holland, J. Hutchin, E. and Kirsh, D. (2000). Distributed Cognition: Towards a new foundation for Human–Computer Interaction research. ACM Transactions on Computer- Human Interaction, 7:2, 174-196.
[6] Nilsson M. and Adams, A. (2005). Building security and trust in online banking. Proceedings of CHI´2005, Portland, USA.

Passive and Active Mediation: Can Conciliation Inform CMC Research?

Matt Billings and Leon Watts
Department of Computer Science, University of Bath, Bath, BA2 7AY
csmmjb@bath.ac.uk; l.watts@bath.ac.uk

There is a body of evidence to suggest that CMC has a distorting effect on the social information used in creating and maintaining relationships. Existing research has tended to identify variables for analysis and used these to infer a general effect of CMC on the majority of relationships. Given the complexity of human interaction, a more relationship-specific approach to the identification of salient variables may be desirable. Conciliation is one such approach: a conciliator must isolate salient factors of a relationship and distort these to transform the relationship that exists between parties. This paper proposes the exploration of the relationship between 'active' (human) mediation and 'passive' (technological) mediation, as a way of informing CMC research. A Model of Relational Communication is proposed, which enables the scope of existing theories of mediation to be evaluated and allows active and passive mediation to be discussed in similar terms. Provisional pilot results indicate that passive mediation impacts on active mediation, and imply that a study of conciliation may inform research into CMC.

CMC, conflict, conciliation

1. INTRODUCTION

Increasing Internet-support for images, video, text and telephony, mean that there are few forms of relationship that cannot be computer-mediated. However, computer-mediation may distort social information to such a degree that relationships are transformed. Existing Computer-Mediated Communication (CMC) research has attempted to identify factors of communication (*e.g.* trust [1], affinity [2], identity cues [3], warmth [4], *etc.*) that can be used as variables for measuring the impact of communication technology on relationships. The cumulative findings of this research, suggest that CMC impacts differently on relationships than Face-to-Face (FtF) communication. However, the measurement variables, identified in these studies, may differ in salience between participants, or change with time. Therefore, the *a priori* selection of variables may fail to appropriately recognise, or fully consider, the salience of other social information present in any given relationship. To fully understand the impact of mediation on relationships, a method for establishing the unique, salient factors of communication within a relationship, must be developed.

Such a method already exists within the field of Alternative Dispute Resolution: conciliation. Third-party conciliation is used to mediate communication between parties in conflict. The conciliator seeks to transform the relationship to one of co-operation[5][6][7]. Social information passed between parties, is distorted in a reflexive manner: the conciliator recognises, analyses and controls the effect they have on social information. Thus conciliation is a form of *active mediation* (AM). This contrasts with the role of technology in CMC. Here, technology distorts social information without reference to context (although parties may consider context when interpreting distorted social information). The technology is not aware of, and therefore cannot control, its effect on the relationship. Thus it is a form of *passive mediation* (PM).

This research attempts to use this distinction between AM and PM to elicit the role of mediation in transforming relationships through the distortion of social information. Investigation of PM will highlight the distorting effect that technological mediation has on social information. Investigation of AM will provide insight as to how salient factors of a relationship can be successfully identified and distorted. An examination of AM conducted using PM, will demonstrate the effect that technology has on salient variables, indicating the effect of CMC on relationships.

2. METHODOLOGY

This study employs both quantitative and qualitative methods. Ecological validity is highly salient when investigating conflict relationships. However, the practical and ethical considerations of requesting parties to participate in computer-mediated conciliation makes large-scale, 'real-world' data collection unfeasible.

Therefore the majority of data will be collected by analysing conflict, engineered under laboratory conditions. Four conditions will be investigated: no-mediation; PM; AM; and AM-PM. However, interim findings will be regularly validated through interviews with, and observations of, conciliation professionals. It is hoped that this will mitigate any negative effects incurred by sacrificing ecological validity for experimental reliability

3. CURRENT FINDINGS

Primarily this study has focussed on establishing the degree to which AM and PM can be discussed in the same terms. This has been undertaken largely through a literature review, resulting in the development of a Model of Relational Communication (MoRC). This MoRC allows both AM and PM to be discussed in the same terms. Therefore, the scope of existing theories of mediation (both passive and active) can be evaluated with this model.

3.1. The Model of Relational Communication

To be able to successfully communicate, parties need to be present at compatible levels of abstraction [8]. Once this is established, present discursive schemas (e.g. shared history, environment, language, technology) will promote certain expectations of behaviour, shaping the way each party is observed by the other [9]. Parties will interpret their observations of the other in terms of the relationship-salient narratives they hold [10][11]. Any form of mediation may unexpectedly distort social information, potentially leading to misunderstanding, or misattribution of meaning. With passive mediation this may be an unintended side-effect of technology-use. With active mediation, this distortion of expectations may be the purpose of the mediation.

3.2. Experimental Findings

Pilot experimental results (from conciliated role-plays) suggest that the introduction of CMC (specifically VMC) into the ADR process has an effect on the conciliator’s methodology and parties’ behaviours. Parties report that VMC promotes feelings of distance and detachment from the interaction, which helps to ‘cool down’ the negotiations. Conciliators using VMC reported feeling more of a ‘passive observer’, necessitating different control and information gathering strategies.

4. CONCLUSION

The MoRC suggests that AM and PM share consistencies, that allow them to be investigated in similar terms. Tentative experimental results demonstrate that AM is affected, to a degree, by PM. Therefore, future research intends to focus on those elements of the conciliator’s strategy that are adjusted to accommodate the introduction of technology into their practice. The reasons for, and result of, these adjustments should provide a useful and unique insight into the effect of CMC on relationship development, maintenance and transformation.

REFERENCES

[1] Olson, G. M., Olson, J. S. (2000). "Distance Matters". Human-Computer Interaction **15** 139-178.

[2] Walther, J. B. Parks, M. R. (2002) Cues Filtered Out, Cues Filtered In: Computer-Mediated Communication and Relationships. In Knapp, M. L. and Daly, J. A. (eds) Handbook of Interpersonal Communication. Sage. 3rd Edition.

[3] Spears, R., Lea, M., Postmes, T. (2000). Social Psychological Theories of Computer-Mediated Communication: Social Pain or Social Gain? In Robinson, W. P., Giles, H. (eds) The Handbook of Language and Social Psychology. Chichester, Wiley.

[4] Short, J. A., Williams, E., Christie, B. (1976). The Social-Psychology of Telecommunications. New York, Wiley.

[5] Deutsch, M. (1987). A Theoretical Perspective on Conflict and Conflict Resolution. In Sandole, D. J. D., Sandole-Staroste, I. (eds). Conflict Management and Problem Solving: interpersonal to international applications. London, Pinter.

[6] Baruch-Bush, R. A., Folger, J.P. (1994). The Promise of Mediation: Responding to Conflict Through Empowerment and Recognition. San Francisco, Jossey-Bass.

[7] Winslade, J., Monk, G, Cotter, A. (1998). "A Narrative Approach to the Practice of Mediation." Negotiation Journal **14**(1) 21 - 41.

[8] Floridi, L. (2005). "The Philosophy of Presence: From Epistemic Failure to Successful Observation." Presence: Teleoperators and Virtual Environments **14**(6).

[9] Goffman, E. (1956) The Presentation of Self in Everyday Life. New York. Doubleday.

[10] McNeil, L. D. (1996). "Homo inventans: The evolution of narrativity." Language & Communication **16**(4). 331-360.

[11] Gover, M. R. (1996). The Narrative Emergence of Identity. Proceedings of the 5th International Conference on Narrative, Kentucky.

Patient Identification and Electronic Healthcare Systems

Valentina Lichtner
Centre for HCI Design, City University, Northampton Square, London EC1V 0HB
at514@soi.city.ac.uk

Patient identification errors are not uncommon. Though, they happen in the complex context of health care provision where million of patients are identified correctly every day - despite their names being imperfect identifiers and their care fragmented among many practitioners. This research investigates how the socio-technical context supports the process of patient identification and how technology can enhance or hinder this process. It entails an ethnographic study of a Walk-In Centre, focusing on how practitioners share among themselves and with other healthcare providers their knowledge of patients' cases and identities. The aim is to contribute both at a theoretical level, in terms of models of 'identity' in electronic systems and related 'identification' processes, and at a practical level, with recommendations and guidelines for the benefit of staff and patients of the Walk-In Centre. The research will also offer a socio-technical view of the study of the concept of 'identity', more usually approached from either technical or individual perspectives.

Identity, healthcare, patient identification, electronic patient records, ethnography, activity theory, distributed cognition

1. PATIENT IDENTITIES, CONTEXT AND ARBITRARY IDENTIFIERS

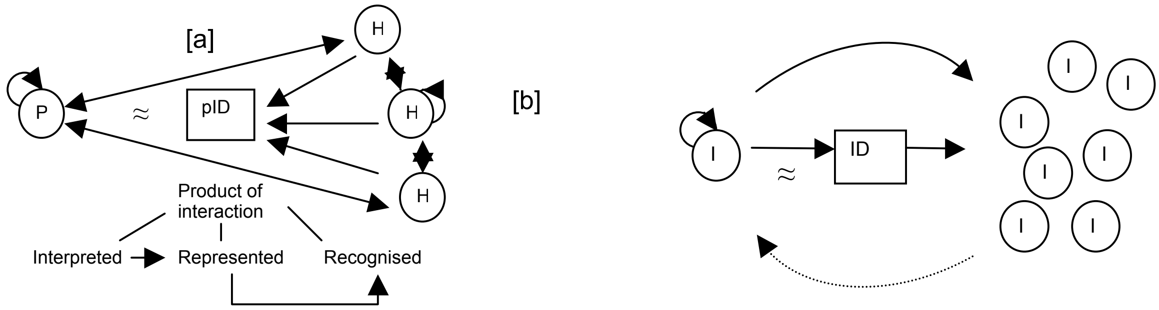
> "She looked up the patient's diet and was trying to order a regular diet. At the fifth screen she saw that the patient was getting tube feeding. This clued her that this was the wrong patient" [field notes, observation of a nurse, U.S. hospital], cited in [1]]"

Patient identification is a safety-critical issue, a priority for the National Patient Safety Agency and all healthcare organisations, and a key factor for the implementation of electronic patient records systems. Electronic patient records (EPR) are often proposed as a solution to a fragmented provision of care and the occurrence of medical errors. But technology can also "foster errors rather than reduce their likelihood" [1] when built on 'utopian' wishes [2] or mistaken assumptions about the nature of healthcare work. This is not based necessarily on single pre-planned tasks performed on computers in isolation but typically situated in complex socio-technical interactions of often interrupted activities in different contexts [1, 3].

Recommendations and guidelines for correct patient identification rely on the redundancy of multiple identifiers. However, in practice, patients are not only identified with their names, date of birth, barcodes, etc., but also on the basis of contextual cues and the narrative of their illness. Names and other identifiers are arbitrary, they do not convey information about the person's identity [4]. In fact, one could say that only the history of the practitioners' interaction with their patients guarantees the correct identification of the patients themselves. Should EPRs support the subjective and personal patient-practitioner experience? Do electronic systems support a shared understanding of patients' identity? This research investigates how the socio-technical context supports the process of patient identification and how technology can enhance or hinder this process.

2. IDENTITY AND TECHNOLOGY

Identity in this context should be analysed as a relational concept. Following an ecological psychology approach, I suggest that '*a patient's identity is what it affords the healthcare practitioner*'. If it is true that the role of a clinician is to keep the patient's trajectory on track [1], then the clinician constructs the 'track' and the patient's identity with the mediation of the patient's record. The practitioner, in the interaction with the patient and with other practitioners, interprets and represents the patient's identity [pID] within the record [Fig.1a]. The patient herself has no control over this representation, except in rare cases. While the identities of patients [P] and healthcare practitioners [H] continuously evolve, the recorded representation [pID] can only ever 'catch-up' in a discontinuous mode. By definition, the recorded representation can never perfectly match the patient's identity and the degree of similarity [≈] between the two may vary. A similar case exists in other identification systems (e.g. identity cards) where the person represented in the system has no control over the representation; this is in contrast with the case of 'social technology' (e.g. chatrooms, etc.), where the individual [I] is more or less in control of the representation of her own personal identity (to be shared with the group) and of the degree of similarity between the two [≈] [Fig.1b].



P = Patient / H = Healthcare practitioner / pID = identity's representation / ≈ = Degree of similarity / I = Individual

FIGURE 1: Relationship individual identity/social contexts mediated by 'healthcare' [a] versus 'social' [b] technology

3. METHOD: BACKGROUND AND LITERATURE RESEARCH, CASES OF ERROR, ETHNOGRAPHIC STUDY

My research to date has been based on an ongoing, multidisciplinary, literature review (in the areas of 'identity', safety critical systems, etc.), supplemented with an understanding of healthcare practices gained from three short visits to hospitals in London, Manchester and Middlesbrough and insight into the 'National Programme for IT' from attending events held by 'Connecting for Health'/Department of Health.

Furthermore, since key aspects of the identification process might become apparent when the process 'fails' - when clinicians get the 'wrong patient' and 'contradictions' may be revealed [5], I have been collecting and analysing cases of misidentification errors reported in worldwide news and elsewhere. Behind the subjective slant typical of some news reporting, this anecdotal evidence suggests that a variety of factors contribute to the (mis)identification process and confirms the idea that misidentification errors are "*an emerging property*" of different kinds of medical activities [4]. Thus, they can only really be understood by studying the whole activity.

The core part of this research will therefore be an ethnographic study of a healthcare setting, studying the process of patient identification in the field over a three month period. The setting will be an NHS Walk-In Centre, a relatively new setting in Primary Care and an area where technology is under-studied as research is more often carried out in hospitals (Secondary Care). At the time of writing the study is scheduled for the end of May 2006. Despite recommendations that technology should be studied in use and in situ, the many barriers to this that require time to overcome make it less frequent than it would be expected or desirable. In the specific case of my PhD, negotiating access to a healthcare setting and achieving NHS ethical approval took more than nine months, bringing the ethnographic study well into the second year of my PhD. The whole process inspired reflective thinking and careful study of qualitative research theory and practice.

4. FUTURE STEPS AND EXPECTED MAIN CONTRIBUTIONS

The research will be informed by activity theory and distributed cognition theoretical approaches. The analysis of the data collected will accompany and follow the ethnographic study. It will be supported by qualitative data analysis software. Through the analysis of the field notes, the intention is to find answers to the research questions and to derive and sustain original theoretical modelling in terms of understanding of 'identity' and 'identification' in healthcare and theoretical tools for the study of technology used in practice. The research will offer a socio-technical view to the literature on 'identity and technology', more usually studied from either technical or individual perspectives. Furthermore, as agreed with the NHS management, the research findings will be reported to the Primary Care Trust and the staff at the Walk-In Centre, hopefully contributing to improving their use of technology.

REFERENCES

[1] Ash, J.S., Berg, M., and Coiera, E. (2004) Some Unintended Consequences of Information Technology in Health Care: The Nature of Patient Care Information System-related Errors. *J Am Med Inform Assoc*, **11,** 104-112.

[2] Gregory, J. (2000) Activity Theory in a "Trading Zone" for Design Research and Practice. In *Proceedings of Doctoral Education in Design: Foundations for the Future, International Conference.* 2000. Staffordshire, UK: Staffordshire University Press, La Clusaz France, 9-12 July 2000 http://heim.ifi.uio.no [10/04/2005]
[3] Berg, M. (1997) *Rationalizing medical work : decision-support techniques and medical practices,* MIT, Cambridge, Mass. ; London.
[4] Sujan, M.A., Henderson, J., and Embrey, D. (2004) *Mismatching between Planned and Actual Treatments in Medicine – Manual Checking Approaches to Prevention*. Human Reliability Associates, for the NPSA. http://www.npsa.nhs.uk [09/09/2005]
[5] Engestroem, Y. (1995) Objects, contradictions and collaboration in medical cognition: an activity-theoretical perspective. *Artificial Intelligence in Medicine*, **7,** 395-412.

Paralinguistic Vocal Control of Interactive Media

Sama'a Al Hashimi
Lansdown Centre for Electronic Arts, Middlesex University, Hertfordshire, United Kingdom
Samaa.alhashimi@gmail.com

The development of multimedia technologies such as augmented and virtual reality has revealed a distinct partiality to visual media [3]. However, sound - and particularly voice- has many dimensions, which have yet to be adequately investigated.

Most developers in the field of vocal input are currently focusing on speech recognition as an alternative input mechanism to keyboards and mice. Although research into speech recognition has made remarkable progress, this technology is still prone to error [1]. Furthermore, the recent prevalence of research into speech recognition has lead many users to the restricted assumption that only the speech aspect of sound can be used to make computers accessible. This supposition has, perhaps, limited many developers' realization of the potential ability of non-speech voice to be used independently as an input mechanism. Consequently, researchers have focused on investigating the verbal aspects of voice rather than on paralinguistic components of vocalization.

One possible outcome of my work is to eventually integrate these paralinguistic components as a complementary input mode to some speech input applications in order to create a synergistic combination that might let the strengths of each mode overcome the weaknesses of the other.

Voice, paralanguage, plotter, voice-physical, vocal telekinesis

1. INTRODUCTION

Vocal utterances can be detected by microphones, and computer software can be used to extract words – and ultimately some aspects of their meaning– from the voice signal. This is the function of speech recognition. However, there are also paralinguistic aspects to vocal utterances – characteristics which may hold information, but without requiring access to the verbal content.

Compared to the difficult process of detecting linguistic cues in speech recognition, which often requires prior training by the user, non-speech features of voice may be near-instantaneously detected through voice signal spectral analysis (e.g., [4]). Speech recognition systems may sometimes be time-consuming especially when they consist of a large list of vocabulary to match with the utterances. On the other hand, the use of paralinguistic vocal control can display a real-time and near-immediate causal relationship between the vocal input and the visual or physical output, and may therefore facilitate continuity and direct engagement. A further distinction of possible interest is that the interpretation of a given signal is usually context-dependent in the case of speech recognition software, whereas in the case of non-verbal input this need not be so. With this in mind, I am investigating how to make creative use of the paralinguistic characteristics of voice in controlling interactive media. I started by developing entertaining implementations of vocal paralanguage, and I inferred that vocal input may permit users to use a wider variety of skills and expressive interaction techniques. Voice as an input to the computer may be an entertaining output to the audience and may potentially act as an engaging sound effect, pushing an interactive installation to performative extremes. Users may become performers who are an integral part of the installation. Moreover, this technique may pave the way for future functional applications which could be used for therapeutic purposes by asthmatic and vocally-disabled users, or as training tools by singers.

2. PROPOSED, CURRENT, AND FUTURE WORK

During the first year, my research encompassed a comprehensive exploration of the various existing voice and speech recognition systems and voice-processing algorithms. I also investigated the range of processes by which sound is detected and represented as well as the range of existing audio-visual applications. This has stimulated my ideas about novel mappings between voice characteristics and visual and physical parameters. Meanwhile, my aim is to focus on the exploration and development of a wider variety of voice-visual mappings which will extend beyond the graphical output to include physical feedback such as changes in the size, temperature, lightness (colour), speed, direction, height, smell, and texture of

real objects. There will be many possible applications for paralinguistic vocal control. Through implementing some of them, I intend to find answers to the following questions:
-What advantages are there in using paralinguistic rather than or along with linguistic input?
-What is the range of applications and advantages of the idea of physical control of inanimate objects with minimal vocal input? And in what sense could this distant vocal control lead to what could be considered a form of *Vocal Telekinesis* by the user?

My first attempt to answer these questions was through developing *SpitSplat*; a game in which voice is used to direct a splat of paint towards moving targets. Pitch moves the splat along the x-axis, and volume moves it along the y-axis. I then developed *SingPong*; a voice-controlled version of 'Pong' where voices and shadows are used to control the height and position of the paddles; volume controls the vertical height of the paddle, and the position of the shadow determines the horizontal position of the paddle.

I also developed *sssSnake*; a two-player *voice-physical* version of the classic 'Snake' game. *sssSnake* consists of an installation table, a coin placed on its surface, and a projected snake. One player utters 'ahh' to move the coin away from the snake, and the other player utters 'ssss' to move the snake towards the coin. The position of players around the table determines the direction of the coin and snake. This encourages physical activity as well as vocal activity. The 'ssss' and 'ahh' are not distinguished through speech recognition but rather through pitch differentiation between the high-pitched 'sss' and the low-pitched 'ahh'.

My next project will be *Blowtter*; a voice-controlled plotter. This plotter will allow the disabled user to blow into a number of closely positioned microphones to control the direction of the plotter head. My future work, will also involve functional implementations of vocal paralanguage through investigating the possibility of using non-speech recognition as a complementary input mechanism to speech-recognition. I already started developing *TOT* (Tip of the Tongue); a wearable device that listens to its user during a conversation. When it recognises vocal segregates such as 'mmm' and 'ahh', it assumes that there is a forgotten word at the 'tip of its user's tongue'. It then listens to the user's attempts to remember similar words and plays an audible list of related words through headphones. When the user utters 'shhh', the audio playback stops. This device might be useful to people suffering from nominal aphasia and who, as a result, have difficulty remembering some words.

3. AREAS OF ORIGINAL CONTRIBUTION TO KNOWLEDGE

I expect the research to contribute novel knowledge to the understanding of how voice can be used on its own and in accord with other input mechanisms, in controlling multimedia applications. I also expect it to provide significant research material for studies about the perception of causality and cognitive relationships in voice-controlled interactive media. By focusing on the paralinguistic applications of voice the work may complement attempts to use speech recognition and other input mechanisms to control interactive applications. Moreover, programming non-verbal voice to move and control inanimate objects has hardly been implemented in any highly developed interactive works. In 1878, Edison built the "phonomotor," which converts voice-induced vibrations, acting on a diaphragm, into motion which can drive a secondary device [2]. Hardly any other highly advanced applications of this technique have been implemented to complement or rival emerging technologies. Furthermore, few studies have created a cross-disciplinary convergence between the field of paralinguistic vocalisations and interactive media. For this reason, I expect the research to contribute in forming new technological approaches towards the association of paralanguage and computing.

REFERENCES

[1] Cox, AL and Walton, A. Evaluating the Viability of Speech Recognition for Mobile Text Entry, *In Proc. HCI 2004: Design for Life.*
[2] Dyer, F.L. (1929) Edison, His Life and Inventions.
(http://www.worldwideschool.org/library/books/hst/biography/Edison/chap10.html).
[3] Gardenfors, D. (2001) Auditory interfaces: A design platform.
http://www.designsounds.net/auditoryinterfaces.pdf.
[4] Quast, H. (2002) Automatic Recognition of Nonverbal Speech: An Approach to Model the Perception of Para-and Extralinguistic Vocal Communication with Neural Networks, Thesis. (University of Göttingen).

Usability and User's Health Issues

Åsa Cajander
Uppsala University, Department of Information Technology / HCI, , Lägerhyddsvägen 2, SE-751 05 Uppsala, Sweden
asa.cajander@it.uu.se

Computer supported work is often stressful and inadequate computer systems and poor usability contribute to the problem. Still the work situation, and work environment of users are seldom considered when developing computer systems. Hence, my research focuses on attitudes about and practices for integrating usability and occupational health issues in IT systems development processes to improve the resulting work situation and well-being of users. The overall goal of the research is to impact software development in practice, hence I do research in real life settings with all the irregularities that occur in such projects.

HCI, action research, user centered system design, usability, work environment, attitudes, perspectives, methods, models

1. INTRODUCTION

Poor usability and a stressful work situation is still a severe problem in computer-supported work, despite efforts to increase focus on these issues. Sweden has a high level of sick rates, particularly in the civil service sector, and some problems relate to inadequate IT systems with poor usability. However, development of computer systems used is largely technology driven, and work organisation and job design are to a large extent shaped by IT systems, since technology often comes before work practices (Clegg, et al, 1997, Eason, 1997). Therefore attitudes about and practices for integrating usability and occupational health issues in IT systems development process are important for the resulting work situation and well-being of users, and here lies my main contribution to the HCI field. Hence my research focuses on attitudes and perspectives in systems development, as embodied in methods and models used for describing users' needs and work, as well as attitudes and perspectives underpinning discourse about users and their work in the development process. What are the attitudes and perspectives, and what are the consequences for the organization and for usability?

2. DESCRIPTION OF RESEARCH PROBLEM

The purpose of my research is to deeper understand perspectives and attitudes in systems development, and how they influence usability of computer systems. I am especially interested in the needs of users characterizing a good work environment, and health aspects of computer supported work, and how long term health goals are perceived in systems development projects. Moreover, I will look at what aspects of human work are illustrated when using traditional software requirements techniques that are based on the idea that complex problems of systems design can be solved using simple check lists and flow diagrams. Is it possible to see how basic assumptions and representations used in systems development affect work situations created by computer systems?

3. THEORETICAL BACKGROUND AND APPROACH CHOSEN

In my research I participate in a large action research project with six different authorities in Sweden. The purpose of the project is to increase knowledge about usability and a good computerized work environment. Moreover, focus is on increasing competence among all parties involved in developing computerized work with lectures, workshops and other information activities.

I position myself as an interpretive researcher, doing qualitative research; however, it is not unlikely that I will do quantitative analysis as a complement in my studies. The interpretive research tradition tries to attain a deeper understanding of reality, and research can be classified as interpretive if it is assumed that our knowledge of reality is gained only through social constructions such a language, consciousness, shared meanings, documents, tools, and other artifacts.

My theoretical framework adheres to situated action (Suchman, 1987), as I perceive users to be highly variable in how they use computers in their work, and often they carry out their work in quite a different way to that modelled or predicted. In my opinion Situated Action and ethnography can provide descriptive accounts of more informal aspects of work, and I see it as a complement to formal methods and models of software engineering.

My research is also based on a constructivist view of language, where we create and understand our reality by using language through communication, and interpretations are flexible, situated and socially constructed.

When people talk about usability they might use the same words on a communication level, but the conversation might still mean different things for them as they have different basic values. Thus basic values load usability with different meanings, assumptions and attitudes. Basic values in this context are deeply rooted ideas that might be known, or unknown to us. Moreover, basic values affect our actions and way of responding to different situations. Hence, basic values can be traced both through communication and action.

Furthermore my theoretical framework partly originates in participatory design which stresses the importance of involving users in the design process and argues that they have a right to be involved in the design of the systems which they will subsequently use. However, there is no clear definition of what research assumptions underpin the approach since the area of PD has been growing rapidly--in terms of numbers of practices, extent of theoretical development, numbers of practitioners, and geographical and institutional diversity of practice. Participatory design has evolved during the last twenty years and has somewhat moved away from its political agenda towards a more pragmatic view where the quality of experience is in focus (Bødker, Ehn et al. 2000).

4. CURRENT STATUS

Our research group has a long tradition in research about work environment and usability in computer systems development. Hence my first paper (Sandblad, 2003) is describing previous research done in this area and it is my starting point, however my contribution in this paper was only as researcher in the project described. In my second paper (Gulliksen, 2003) we focused on communicating an approach and an attitude to user-centred design through key principles. In this paper I took active part in the discussion, and the key principles reflect my previous experiences of systems development. In my third paper (Cajander, 2006 a) we have looked at management and their perspectives on usability in a public authority in trying to understand the problem of poor usability in computer-supported work. What are their interpretations of usability and their view of usability? Why do managers interpret usability as they do, and what are the consequences for the organization and for usability? My last publication (Cajander, 2006 b) describes a large interview study conducted in six authorities in Sweden. In this study we have tried to see attitudes and perspectives underpinning discourse about users, usability and work and discuss how these perspectives affect usability work in these organisations.

5. FUTURE WORK

My future research includes more extended field studies and participation in software development projects at the authorities included in our research project described above. Here I will learn more about case study research, the role of the researcher in action research projects as well as writing up qualitative research. Moreover, we will also work with quantitative measurements of usability, and it will be interesting to use see how this affects the organisation.

REFERENCES

Bødker, S., P. Ehn, et al. (2000). Co-operative Design - perspectives on 20 years with "The Scandinavian IT Design Model". NordiCHI, Stockholm.

Cajander, Å., Boivie, I., Gulliksen, J. (2006 b) Usability and User's Health Issues in Systems Development - Attitudes and Perspectives. In E. Law, E. Hvannberg, G. Cockton (eds.) Maturing Usability; Quality in Software, Interaction and Value. Springer Verlag.

Cajander, Å., Gulliksen, J.,Boivie,I. (2006 a) Management Perspectives on Usability in a Public Authority – a Case Study. Submitted to NordiCHI 2006

Clegg, C., Axtell, C., Damodaran, L., Farbey, B., Hull, R., Lloyd-Jones, R., Nicholls, J., Sell, R. And Tomlinson, C. (1997) Information technology: a study of performance and the role of human and organisational factors. Ergonomics, 40(9), 851 – 871. Cockton (eds.) Maturing Usability; Quality in Software, Interaction and Value. Springer Verlag

Eason, K., (1997). Understanding the Organisational Ramifications of Im-plementing Information Technology Systems, in Helander, M. Landauer, T.K. Prabhu, P. Handbook of Human-Computer Interaction, Second revi-sion, Ch 61; pp 1475-1494; Elsevier Science B.V., Amsterdam, the Nether-lands.

Gulliksen, J., Göransson, B., Boivie, I., Persson, J., Blomkvist, S. & Cajander, Å. (2003) Key principles for user-centered systems design. Behaviour and Information Technology, Vol. 22, No. 6. pp.397-409.

Sandblad, B., Gulliksen, J., Åborg, C., Boivie, I., Persson, J., Göransson, B., Kavathatzopoulos, I., Blomkvist, S. and Cajander, Å. (2003) Work environment and computer systems development. Behaviour & Information Technology, Vol. 22, No. 6, pp. 375–387

Suchman, L.A. (1987) Plans and Situated Actions: The Problem of Human Machine Communication. Cambridge University Press. Cambridge, Ma, USA.

Ambient Intelligent Environments: an Ambient Intelligent Navigation-Assistance System for the Visually Impaired

Luke Okelo, David England, A. Taleb-Bendiab
School of Computing and Mathematical Sciences,, John Moores University
James Parsons Building, Byrom Street, Liverpool L3 3AF, United Kingdom
{cmplokel@livjm.ac.uk, D.England@ljmu.ac.uk, a.talebbendiab@ljmu.ac.uk}

We describe our progress on the development of a prototype for an ambient intelligent navigation system for the blind. The purpose of such a navigation system would be to enable blind and visually impaired individuals to navigate easily within both familiar and unfamiliar environments in a building or within a city.

We envisage that such a system would operate over a distributed wireless network, and control the ubiquitous devices embedded within the environment to offer navigation-assistance for the visually impaired users.

Ambient Intelligence, blind navigation systems, visually impaired, blind users

1. BLIND NAVIGATION SYSTEMS

Whereas computer technology has made it possible, often quite straightforward and inexpensive to make words accessible to blind people, making anything except words accessible remains a formidable challenge [1]. Technology that has been adapted for the blind is known as Typhlotechnology. Typhlotechnology tackles the objective of developing new technologies-or of adapting and providing accessibility to existing technologies-for use and utilization by people with blindness or visual impairment so that their experience can be as similar as possible to those who without disabilities. The visually impaired are at a considerable disadvantage when it comes to navigation, for they often lack the type of information that is available and essential to sighted individuals who navigate. Initial technology developments were limited to devices that helped individuals avoid obstacles i. e. Electronic Travel Aids [2]. Electronic Travel Aids (ETA) are a form of assistive technology having the purpose of enhancing mobility for the blind pedestrian. The most commonly known device is the laser cane. Other examples of electronic travel aids include e.g. the long cane, ultrasonic obstacle avoiders, tactile maps, spoken directions etc. ETA's were not fully beneficial in that they did not promote independence in the mobility of blind or partial sighted persons, nor did they facilitate their successful navigation within unfamiliar environments. They were often inconvenient to port and could only be customized manually with great difficulty. Most blind users do not find the slight improvement in mobility performance and the additional worry of maintaining a complex, expensive battery operated system that must be carried around and kept track of to be worth the extra cost. Global positioning systems for the visually impaired was first introduced in late 1980s.There are now a number of research and commercial endeavours around the world utilizing GPS or DGPS for determining the position of a blind traveller. There have been many attempts made at integrating GPS into navigation assistance systems for the blind e.g. MoBIC, Drishti, NOPPA, BrailleNote GPS and Trekker. However, now that the technology is portable, it still faces the challenge of lacking in awareness and understanding of the value of location information

1.1. Ambient Intelligence

Ambient intelligence can be defined as "a pervasive and unobtrusive intelligence in the surrounding environment supporting the activities and interactions of the users" [3]. The objective of ambient intelligence is to "broaden the interaction between human beings and digital information technology through the usage of ubiquitous computing devices [4]. Key technological requirements for Ambient intelligence are: - very unobtrusive hardware, a seamless mobile/fixed communication infrastructure, dynamic and massively distributed device networks, natural feeling human interfaces and dependability and security" [5]. Therefore, ambient intelligence implies a seamless environment of computing, advanced networking technology and specific interfaces. An ambient intelligent blind navigation system would operate over a distributed wireless network to control the ubiquitous devices embedded within the environment and provide navigation-assistance for the visually impaired users. Operating over a distributed wireless network would ensure that the system will surround the user; the ubiquitous devices would be embedded in the environment to ensure that they provide navigation-assistance in a non-intrusive way. In order to be sensitive to the individual users, the ambient intelligent system would need to recognise, through the use of sensors and actuators, any individual users of the system. The system would also need to be able to receive natural voice orders from the individual blind users, by having a multi-modal, voice based human interface through which the users could relate to it.

2. FUTURE WORK

The prototype currently under design represents a blind navigation system in an ambient intelligence framework. Whereas most systems that have been developed so far for the visually impaired still face the challenge of lacking in awareness and understanding of the value of location information, the ambient intelligent blind navigation system, by integrating several technologies including sensors and actuators, distributed wireless networks, agent technology and ubiquitous, embedded computing devices, provides navigation-assistance to the blind. However, the seamless environment of computing, advanced networking and specific interface will also enable the provision of context-aware services as one way of exploiting the potential of ambient intelligence. The next step is to ensure that the navigation-assistance offered by the system to the user is based on dynamic data, e.g. will the ambient intelligent blind navigation system act as a supplementary aid to other current existing blind navigational aids or stand-alone on its own?

REFERENCES

[1] http://www.once.es/otros/premios/imasd/docs/dprensa-i.doc
[2] http://www.noogenesis.com/ETA.html
[3] Journal Article: Riva G, Vatalaro F, Davide F, Alcaniz M, (2005) “The psychology of Ambient Intelligence: Activity, Situation and Presence” www.emergingcommunication.com, 6, 17-33
[4] Journal Article: Riva G, Vatalaro F, Davide F, Alcaniz M, (2005) “New Technologies For Ambient Intelligence”, www.emergingcommunication.com, 6, 3-15
Conference:
[5] Lindwer Menno, Marcules Diana, Basten Twan, Zimmermann Rainer, Marculescu Radu, Jung Stefan, Cantore Eugenio, (2003) “Ambient Intelligence Vision And Achievements: Linking Abstract Ideas To Real-World Concepts”, Proceedings of DAT ’03, Europe, March 03-07, http://www.ece.cmu.edu/~sid/pubs/papers/DATE03.pdf

Fidelity Requirements of a Human Factors Research Train Driver Simulator

Thomas K. Yates
Virtual Reality Applications Research Team (VIRART), Human Factors Research Group,
School of Mechanical, Materials and Manufacturing Engineering, The University of Nottingham,
University Park, Nottingham, NG7 2RD, UK
eaxtky@nottingham.ac.uk

1. INTRODUCTION

Network Rail is responsible for maintaining and renewing the rail infrastructure in the United Kingdom, ultimately it will be train drivers who will be affected by any changes. Therefore, Network Rail requires some tool to determine the impact of the infrastructure changes on the driver, essentially to assess the driveability and safety of routes post infrastructure change. Within the rail industry, there are a number of areas of human factors being researched and there is a need for some of this research to be undertaken in a simulator. There is no rail simulator in the United Kingdom which is specifically for research purposes; human factors researchers working in the rail industry have reported that having such a simulator would be very useful, in some projects essential, when undertaking human factors research.

However, there are still some unresolved issues regarding simulation, particularly with regard to how realistic the simulators have to be in order to produce valid results from experiments. So before any designing or building of a simulator commences there are many factors to consider with regard to its design. One particular area of interest is the absence of guidelines for the fidelity requirements. As such, this research hopes to address the subject of simulation fidelity in the area of train driver simulators for human factors research.

2. METHODOLOGY

The overall process of the PhD will be completed by undertaking three studies. Study One consisted of Subject Matter Expert interviews, these were used to determine what human factors specialists consider important to include in a simulator design. The data from these interviews helped define a User Requirements Specification for a simulator where the users are the human factors researchers. The follow up to this, Study Two, will be the development of virtual environments of varying fidelity. This will be for testing what the Subject Matter Experts stated were their requirements for the simulator and also to assess the fidelity requirements for the train driving task. Then the final part to the research will be Study Three, which will be the establishment and implementation of a validation technique for the developed simulators in order to assess the validity of results obtained from a study. In order to complete these studies it is intended that a range of virtual environments will be developed.

3. SIMULATOR DEVELOPMENT AND FIDELITY WORK

The User Requirements Specification has been used to guide the development of a set of train simulator virtual environments. Study One identified a lack of guidelines with regard to what information should be included in the visual scene, as well as a lack of certainty regarding the quality level required for different types of information within a simulation.

Study Two proposes to determine this required quality level by developing fidelity guidelines for use in train driver simulator design. There are no accepted fidelity metrics for simulated environments, and a review of the literature shows there are many different techniques which could be adopted for determining the appropriate level of fidelity in a virtual environment. A technique that is to be adopted is one proposed by Watson [1[which combines naming times, ratings and preferences techniques to measure the level of fidelity in a simulation. This technique would seem appropriate if, for the naming times technique, the rail industry's rules on minimum sighting times of lineside signals and indicators are incorporated. If this fidelity metric methodology is combined with the railways standards on minimum sighting times then it is possible to use this as the basis for an experiment to determine the required level of fidelity in a train driver simulated environment.

A dynamic environment will be presented to the driver, this will consist of models of trackside items of differing levels of fidelity. If the driver can identify the item within a predefined time, as stated in the industry standards, then that will be identified as the appropriate level of fidelity which will ensure the driver is able to drive as he would in the equivalent real world situation. An example of the types of environment that are under development are shown below in Figure 6, the image on the left shows an example of a virtual environment that consists exclusively of 2-dimensional models, whereas the image on the right consists of 3-dimensional models. The experiments will assess the impact of these differing levels of fidelity and results from this study will be presented at the Doctoral Consortium.

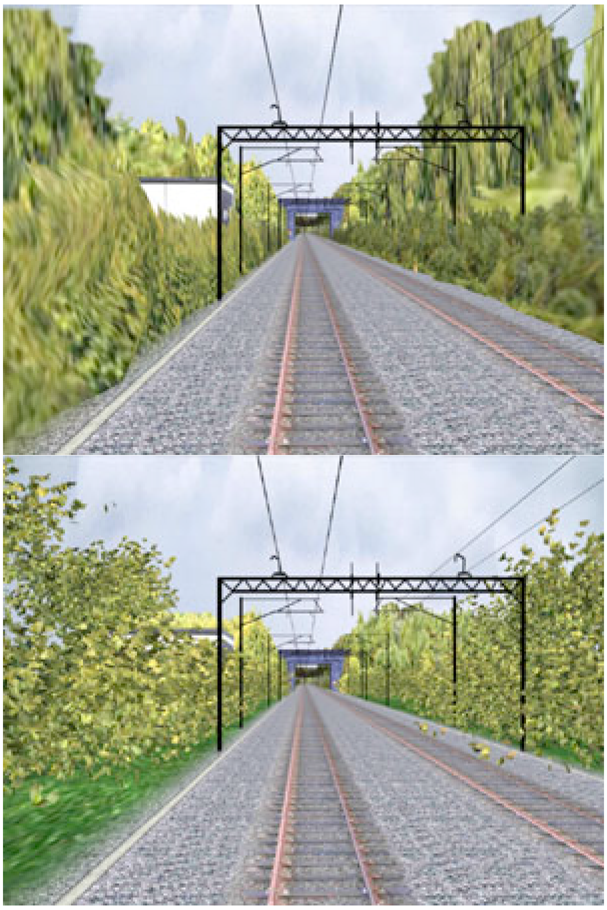


Figure 6: Virtual environments consisting of 2-dimensional and 3-dimensional models

4. CONCLUSIONS

A domain review has been completed and the area of research for this PhD has been identified as the development of fidelity guidelines for a train driver simulator for human factors research. To date there has been no published work relating to the requirements for a research simulator, Study One addresses this situation and with the assistance of a cross section of the human factors community a list of simulator requirements have been developed. The next stage of the research is to concentrate on designing a set of experiments to help determine what is the level of fidelity required in a train driver simulator in order to ensure that the train driver still drives as he would in the real world situation. This will be undertaken by designing a set of virtual environments where there are varying levels of fidelity in each, if the driver can recognise and react to the infrastructure within a stated time, the time denoted in the industry guidelines, then the simulated environment can be considered to be of sufficient quality.

REFERENCES

[1] Watson, B., Friedman, A. and McGaffey, A. (2001) Measuring and predicting visual fidelity. Proceedings of the 28th annual conference on Computer graphics and interactive techniques, pp. 213-220. ACM Press, New York, USA.

An Automatic Web User Attention Analyser

Abdallah Namoune
School Of Informatics, University of Manchester,
PO Box 88, Sackville Street, Manchester, M60 1QD, UK
Abdellah.namoune@postgrad.manchester.ac.uk

In the last decade, eye tracking has been increasingly used to study user eye movements in web pages. Despite its efficiency and accuracy, eye tracking is considered monotonous and time consuming as it requires extensive data reduction. This research proposes building an Automatic Web User Attention Analyser. It is a predictive system intended to simulate human visual attention while browsing web pages. Constructing such a promising tool will enable designers to early test their hypotheses regarding the complex interaction between user eye movements, user goals, and visual elements of the web (animation, images, headers…etc). Additionally, this predictive tool will suggest design advice to improve the representation of web pages.

Visual attention, eye movements, scanpath, saliency-driven, goal-driven

1. INTRODUCTION

Web designers continuously attempt to understand the user's ocular behaviour while viewing different web pages in order to establish a cognitive basis for their interface design. Some of the interesting questions which designers try to answer are: Where do users look at when they browser web pages? Do they notice critical elements? One approach to solving these questions is to perform eye tracking. However, a more cost effective approach necessitates building a configurable predictive tool which simulates user eye movements.

Unlike previous models [1, 2] which mainly incorporate a semantic component to predict user navigation behaviour, our model will asses the saliency-driven factors (animations, images, size…) and predict user fixations while viewing web pages. The Automatic Web User Attention Analyser is based on attention factors that have been synthesised from attention theories [3, 4] and previous eye tracking studies [5].

This research attempts to fulfil the following objectives: (a) Build a computational predictive tool which simulates user eye movements, and outputs a prediction of the visual scanpath which users follow when they browse web pages. (b) Investigate the influence of visual attention on user comprehension and user memory of web pages.

2. METHODS

2.1 System Development

This stage involves implementing the components of the system, as summarised in figure 1.

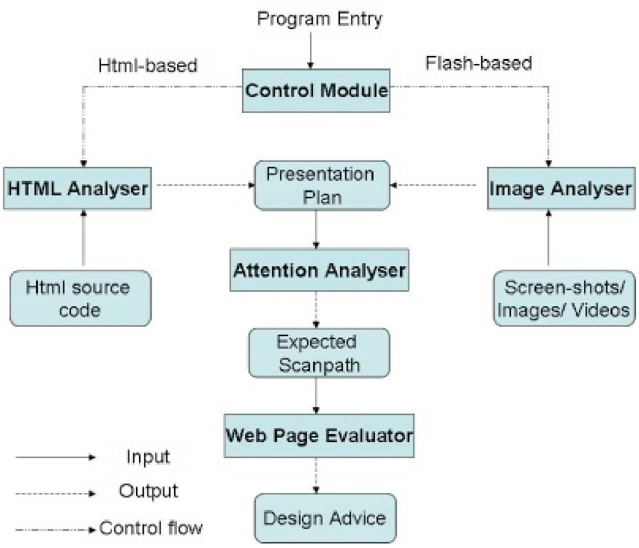


FIGURE 1: Components of the Automatic Web User Attention Analyser

The algorithms that will be used in the tool are as follows:

- Web page segmenter: this algorithm is responsible for segmenting HTML-based web pages into visual blocks.
- Image segmenter: this algorithm segments flash-based web pages into coherent visual blocks.
- Change (video) detector: this algorithm detects animation.
- Saliency calculator: this algorithm uses different visual attributes (motion, size, position, contrast, brightness, and media type) to calculate and assign an appropriate saliency value to each visual block of the presentation plan.
- Semantic assessor: this algorithm will be used to assess the similarity between user goals and the content of the web pages.
- Nearest-neighbour search: this algorithm finds the closest salient visual block to the current focal point.
- View-order predictor: this algorithm executes the scanpath in a left to right, top to bottom approach ignoring the effect of saliency.

The attention factors used to generate an expected scanpath are briefly summarized in table 1:

Salience-driven Factors	**Goal-driven Factors**
Context Evaluation (Contrast, Peripheral vision)	Semantic similarity to user goals
Saliency (Motion, Size, Colour, Graphical Elements, Text Style)	
Read Order Sequence (Spatial information)	

TABLE 1: Attention Factors of the Automatic Web User Attention Analyser

2.2 Eye Tracking

A controlled user-study will be conducted to collect eye movements in order to evaluate the validity of the model. If the expected scanpaths of the Automatic Web User Attention Analyser and the real scanpaths obtained from users are different, this information will be used to inform the attention factors of our model.

3. CONTRIBUTIONS

Web pages contain a mixture of textual, pictorial, and multimedia content. Having many competing elements for attention in a web page can be ineffective as it could delay or prevent delivering the appropriate information to users. Therefore, the Automatic Web User Attention Analyser will provide advice to web designers regarding the presentation of information, and enable them to establish a good flow of attention according to the importance of items within the content. This research will also contribute towards a better understanding of the interaction between saliency-driven factors and goal-driven factors. Finally, the introduction of memory tests will investigate the relationship between visual attention and user memory.

4. CURRENT STATUS AND FUTURE WORK

In the first year of my PhD which I started in October 2004, an extensive literature review related to visual attention and eye movements on the web has been carried out. The requirements specification of the system has been completed in the first year. At the time of writing, the HTML Analyser and the saliency-driven factors of the Attention Analyser module have been integrated in the system prototype.

A user experiment of 30 subjects is currently being conducted to test the validity of the system. Data collection and analysis is expected to be completed by July 2006. The final part of the prototype will consider implementing the Image Analyser and including the critique for the web pages.

REFERENCES

[1] Chi, E.H., Pirolli, P., Chen, K., and Pitkow, J. (2001) Using Information Scent to Model User Information Needs and Actions on the Web. Proceedings of the SIGCHI conference on Human factors in computing systems, Seattle, Washington, United States, pp.~490-497. ACM Press New York, NY, USA.

[2] Kitajima, M., Blackmon, M.H., and Polson, P.G. (2000) A Comprehension-based Model of Web Navigation and Its Application to Web Usability Analysis. People and Computers XIV - Usability or Else! (Proceedings of HCI 2000), pp.~357-373. Springer.
[3] Palmer, S.E. (1999) Vision Science: Photons to Phenomenology. MIT Press, Cambridge, London.
[4] Itti, L., Koch, C., and Niebur, E.(1998) A Model of Saliency-Based Visual Attention for Rapid Scene Analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20, 1254-1259.
[5] Outing, S. and Ruel, L. (2004) The Best of Eyetrack III: What We Saw When We Looked Through Their Eyes. Retrieved 25/05/2005 from http://www.poynterextra.org/eyetrack2004/main.htm

Studying Emotions and Non-Instrumental Qualities as Parts of the User Experience

Sascha Mahlke
Centre of Human-Machine-Systems, Berlin University of Technology, Jebensstr. 1 – J2-2, 10623 Berlin, Germany
sascha.mahlke@zmms.tu-berlin.de

Designing the user experience has become a central focus of interactive system developments in recent years. Several attempts have been made to broaden the traditional focus on the efficient achievement of goals and incorporate a fuller understanding of additional aspects of the user experience. The current research project investigates the role of non-instrumental quality aspects like hedonic and aesthetics qualities and their interplay with emotions in shaping the user experience. An integrative model, the methodological approach and first empirical results are presented. Future research will combine methods for measuring emotional aspects and relevant quality dimensions to further understand the interplay of various aspects of the user experience of interaction.

User experience, emotional user reactions, aesthetics, hedonics, usability, evaluation

1. INTRODUCTION

Definitions of usability focus on tasks and goals, their efficient achievement, and the cognitive information processing involved. To go beyond these traditional perspectives and for a better understanding of how people experience technology, various approaches have been suggested that take other aspects of the interaction into consideration. Non-instrumental quality aspects and the role of emotions are discussed as two important areas for research on the user experience [1]. Jordan [2] argued for a hierarchical organization of user needs and claimed that along with the functionality and usability of the product, different aspects of pleasure are important to enhance the user's interaction with it. Further analyses studied selected non-instrumental quality aspects of interactive systems in detail, such as hedonic quality [3] and visual aesthetics [4]. Recently, the term emotional design [5] has received significant attention. Desmet & Hekkert [6] went a step further by presenting an explicit model of emotions according to product perceptions. Zhang & Li [7] studied the concept of affective quality as the ability of interactive systems to cause changes in the user's affective state. In this way non-instrumental quality aspects and the role of emotions were studied individually for a more in-depth understanding. However, to assess interactive systems regarding the user experience as a whole these various aspects have to be integrated to fully understand and compare users' experiences of interaction with different systems. Rafaeli & Vilnai-Yavetz [8] studied the interrelations between instrumental and non-instrumental quality aspects as well as emotions in a non-interactive domain. Tractinsky & Zmiri [9] transferred this approach to the area of websites. My research aims to carry on this first steps and wants to lay a more elaborate theoretical basis, use a broader methodological approach and provide further empirical results on this research problem.

2. THE USER EXPERIENCE RESEARCH FRAMEWORK

The user experience research framework presented in Figure 1 integrates the discussed aspects of the user experience and provides the basis for further research [10]. Instrumental and non-instrumental quality aspects are summarized in two distinct components that include the various quality dimensions described earlier. Properties of the interactive system that are perceived by the user while interacting with the system influence how the user experiences the product on these quality dimensions. On the other hand, these quality perceptions influence user's behaviour and judgments. Emotional user reactions play another important role.

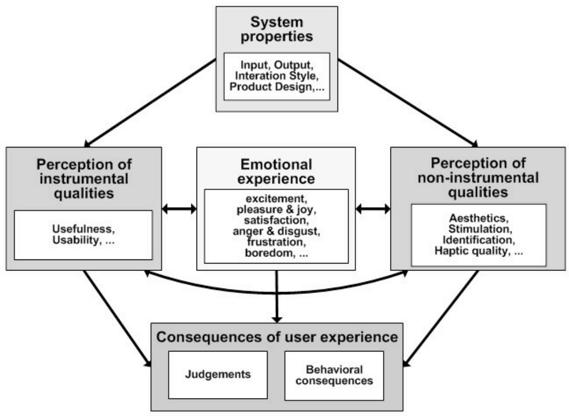


FIGURE 1: User experience research framework

They are influenced by quality perceptions and also have an influence on consequences of the user experience. In the current research I focus especially on the interplay of instrumental and non-instrumental quality aspects with emotions.

3. METHODOLOGICAL ASPECTS

Rafaeli & Vilnai-Yavetz [8] took a qualitative approach to this research problem. Besides, a lot of questionnaires were developed to assess instrumental and non-instrumental quality aspects [3, 4, 6, 7]. These can be used to survey various experience dimensions. The methodological focus of my research lies on new methods to measure emotional reaction as parts of the user experience. We conducted a study to compare and integrate different approaches to emotion measurement. Based on a multi-component approach to emotions [11] different aspects of emotions in an interactive context were investigated: subjective feelings, physiological activation, motor expressions, cognitive appraisals, and behavioural tendencies [12]. The results suggest that a combination of methods that assess different components of emotional reactions provide a comprehensive basis for analyzing emotions in human-technology interaction.

4. EMPIRICAL RESEARCH

Furthermore, a first explorative study was conducted using the research framework as the basis for assessing user experiences with interactive systems [13]. Four digital audio players were chosen for the study. All were from the same manufacturer, so we did not have to deal with the influence of brand. Nonetheless, players differed in terms of various design aspects. Thirty individuals (fifteen women and fifteen men) participated in the study. They were between 20 and 30 years old, most of them students at Berlin University of Technology. All participants tested each product. Four short tasks were given to the participants for each product. After accomplishing the tasks, participants filled out a questionnaire that assessed ratings on different experience dimensions (usefulness, ease of use, visual aesthetics, hedonic qualities, physio-pleasure) and emotional consequences. After using each of the players, participants made a ranking list of the players. The preliminary results give first hints on the complex interplay of instrumental and non-instrumental quality perceptions with emotional user reactions. Instrumental quality aspects, i. e. the usability of the system have a main influence on the emotional user reactions, but also the non-instrumental aspects play a significant role. Further studies will focus on the role of quality perceptions for specific emotional reactions like satisfaction, enjoyment, anger or unhappiness.

5. CURRENT STATUS AND NEXT STEPS

The research framework serves as basis for the next steps. Furthermore, the results on the use of different methods to assess emotional user reactions will be applied in further empirical studies and the explorative study gave first hints on the interrelations between emotions, instrumental and non-instrumental quality aspects in human-technology interaction. In two following experiments these interrelations will be studied in more depth. Therefore, prototypes of interactive products will be designed and varied on various dimension that have an influence on specific instrumental, e.g. usability, and non-instrumental, e.g. visual aesthetics, quality perceptions. Thus, it should be possible to better understand in which way differences in design properties influence quality perceptions that determine emotional experiences as well as overall judgements and behavioural consequences.

ACKNOWLEDGEMENTS

This research is supported by the German Research Foundation (DFG) as part of the Research Training Group 'Prospective engineering of Human-Technology Interaction' (no. 1013).

REFERENCES

[1] Hassenzahl, M. and Tractinsky, N. (2006). User experience - a research agenda. Behaviour & Information Technology, **25**, 91-97.
[2] Jordan, P. W. (2000). Designing pleasurable products. Taylor & Francis, London.
[3] Hassenzahl, M. (2001). The effect of perceived hedonic quality on product appealingness. International Journal of Human-Computer Interaction, **13**, 481-499.
[4] Lavie, T. and Tractinsky, N. (2004). Assessing dimensions of perceived visual aesthetics of web sites. International Journal of Human-Computer Studies, **60**, 269-298.
[5] Norman, D. A. (2004). Emotional Design: why we love (or hate) everyday things. Basic Books, New York.
[6] Desmet, P. and Hekkert, P. (2002). The basis of product emotions. In: W. Green and P. Jordan (eds.), Pleasure with Products, beyond usability. Taylor & Francis, London.
[7] Zhang, P. and Li, N. (2005). The importance of affective quality. Communications of the ACM, **48** (9), 105-108.

[8] Rafaeli, A. and Vilnai-Yavetz, I. (2004). Instrumentality, aesthetics and symbolism of physical artifacts as triggers of emotion. Theoretical Issues in Ergonomics Science, **5**, 91-112.
[9] Tractinsky, N. and Zmiri, D. (2006). Exploring Attributes of Skins as Potential Antecedents of Emotion in HCI. In P. Fishwick (ed.), Aesthetic Computing. MIT Press, Cambridge, MA.
[10] Mahlke, S. (2005). Understanding users' experience of interaction. Proceeding of EACE '05, Chania, GR, 29.09.-01.10., pp. 243-246, National Technical University, Athens.
[11] Scherer, K. R. (1984). On the nature and function of emotion: A component process approach. In K. R. Scherer & P. Ekman (eds.), Approaches to emotion. Erlbaum, Hillsdale, NJ.
[12] Mahlke, S., Minge, M. and Thüring, M. (2006). Measuring multiple components of emotions in interactive contexts. Extended Abstracts of CHI '06, Montreal, CA, 22.-27.04., pp. 1061-1066, ACM Press, New York.
[13] Mahlke, S. (2006). Studying user experience with digital audio players. Proceedings of ICEC' 06, Cambridge, UK, 20.-22.09., Springer, Berlin.

Genome Visualisation

Joanna Jakubowska
University of Glasgow, Department of Computing Science, Glasgow G12 8QQ, Scotland
asia@dcs.gla.ac.uk

In the visualisation of complex data there are a lot of unsolved problems. The complexity of data is increasing fast, but the users' ability to understand seems to be constant. The biological researchers we collaborate with would like to see the data in some user-friendly tools, but, the tools, computer monitors, and machines have their limitations and do not always precisely show the data under investigation.

We experiment with visualisation tools and develop new techniques which will more precisely express the complexity of data. We examined existing tools used by researchers and created a classification which will help to find a solution. We developed a prototype called VisGenome. VisGenome is an application for visualising single and comparative representations of the rat, the mouse, and the human chromosomes. During my PhD, I am going to conduct a user study and experiments with both existing and new applications for genome visualisation.

Genomics visualisation, genome browser

1. INTRODUCTION

Large and complex data give rise to visualisation problems. Researchers try to solve the problems either from database or visualisation point of view. There is only a small group of people examining current work practices and studying Human-Computer Interaction with biological data. We began our work by studying existing visualisation solutions in order to find out what features they offer, which of those correctly support data analysis, and which are not helpful. The study will allow us to find a better solution for data analysis which overcomes cognitive problems.

In the area of genome visualisation, we found two groups of problems: visualisation and management of large amounts of data which should be shown at the same time. We would like to find a solution which will clearly present the information, including all relevant data the biologists wish to see. We are aiming to derive general principles of data representation and visualisation usability for genomics. We also would like to discover how best to compare data coming from various sources and experiments in biological setting.

The PhD work focuses on the use of visualisation to support the understanding of very large data sets. We would like to create an universal solution. We hope, VisGenome will solve both the visualisation problems and some of the data integration problems. We would like to offer a clear presentation of the data the biologists wish to see. We cooperate and make experiments with biologists where we study Human-Computer Interaction. We hope that our study will allow for deeper understanding of the problems and help in finding a solution.

2. PROBLEMS IDENTIFIED DURING RESEARCH

There are two main groups of problems, one related to visualisation and the other one to database integration.

We found out that there is no universal tool for genome visualisation. Each biologist group uses a different tool in their experiments. Some tools show all publicly available data, see Ensembl [3]. Other tools visualise only specific data from one experiment and provide no possibility to add any external data. There is no possibility to add data from a new experiment and compare the results. The biological data has a variety of formats and is situated in a lot of places. Genome browsers can read the data from special file formats or from a database. Frequently, the databases use different technologies and the users need to convert the data before it can be visualised.

The main visualisation problem is how to show all data and what kind of visualisation technique to use. The tools offer poorly designed zooming or panning. Some applications, for example DerBrowser [2] offers zooming, but it is not smooth zooming and is limited in depth. New genome browsers are often designed without respect for the principles of HCI. Therefore we decided to experiment with VisGenome. At the time, the application shows data in natural scaling, and we will extend it to use cartoon scaling [1]. We are developing an algorithm which represents different kinds of data in cartoon scaling.

3. PROPOSED SOLUTION / METHODOLOGY

The methodology will be based on prototyping with the users, combined with investigating existing genome browsers. In cooperation with the biologist group we tested some more popular genome browsers in order to find which one supports the interpretation of their experiments. We found out that none of the browsers fully supports the user requirements we identified. The experiments motivate us to define a genome browser classification and develop a new visualisation tool - VisGenome. We are going to conduct a study with the users – biologists from schizophrenia and hypertension groups at the University of Glasgow, London and Edinburgh. The experiments will combine studying data from the biologists' experiments with measuring the time, and counting the mouse clicks. A survey will be used to get users impression on the legibility of the display, aesthetic appeal, and the subjective ease of use.

4. CONTRIBUTIONS

The experimental work will be the main contribution. To our knowledge, nobody has carried out so far experiments with genome browsers. The deeper understanding of how the biologists work, what kind of information they need in their experiments, and the two-fold character of my research (the experiments are parallel with the prototyping of VisGenome and looking for an universal solution for representation the relevant biological data) is the next contribution.

5. CURRENT WORK

Initially, we studied the existing source code of two browsers SyntenyVista [1] and DerBrowser [2]. The aim of the work was to modify the existing implementation to implement Fisheye [5] and Excentric Labelling [6]. Then, we surveyed genomics visualisation software and defined a classification of genome browsers according to three dimensions: number of species, size of the objects shown, and representation complexity. The classification argues the need for a new genome browser which offers improved zooming functions. Therefore we develop an extension of SyntenyVista [1], VisGenome. The application allows for the addition of new data types to the display, and will be able to fully satisfy user requirements. So far, we developed the part responsible for single data representation. The tool offers smooth zooming and panning implemented using Piccolo [4]. The users can keep an area of interest in focus during zooming process. The solution allows the biologists to keep the context which help them not to get lost.

6. FUTURE DIRECTIONS AND CONCLUSION

Future work will use a number of prototypes which will be assessed with users. We will test not only VisGenome but also other tools with users.

Visualisation of genome comparisons is an important research tool in biology and medicine. There are variety of genome browsers which in practice should perform the same function – show the chromosomes of some species in detail. The differences in the view and also in functionality of genome browsing motivated me to create a classification of genome browsers and, in consequence, to develop new tool – VisGenome. We believe, that biologists still require new methods to visualise genomic data.

REFERENCES

[1] Hunt, E. et al. (2004) The visual language of synteny. *OMICS*, **8(4)**, 289–305.
[2] Leser, U. et. al. (1997) IXDB, an X chromosome integrated database. *NAR*, **26(1)**, 108-111.
[3] Ensembl, http://www.ensembl.org.
[4] Piccolo Toolkit, http://www.cs.umd.edu/hcil/piccolo/.
[5] Furnas, G. W. (1986) Generalized Fisheye Views. *Human Factors in Computing Systems CHI '86 Conference Proceedings*, 16–23.
[6] Fekete, J. D. and Plaisant, C. (1999) Excentric Labeling: Dynamic Neighbourhood Labeling for Data Visualization. *CHI'99 Pittsburgh PA, USA*, 512–519.

Design and Evaluation of Tangible Interfaces for Children

Diana Xu
Child-Computer Interaction (ChiCI) Group, Department of Computing
University of Central Lancashire, Preston, PR1 7BU
yfxu@uclan.ac.uk

My research is on the design and evaluation of technologies for children that go beyond the desktop computer and which increasingly merge the physical and digital worlds, especially Tangible Technologies. We look at the traditional evaluation methods for children's interactive products and in search for the methods for children's Tangible Technology. The research will contribute to paradigms such as the design and evaluation of 'disappearing computer' and 'tangible computing'. The research involves empirical evaluation studies and the design of prototype systems.

Tangible, children, design, evaluation

1. PROBLEM STATEMENT

Tangible technology has great educational potential [6]: unlike traditional desktop computers with GUI interfaces that rely on a limited variety of physical objects and a limited range of our abilities, Tangible User Interfaces (TUIs) offer direct manipulation of physical objects and provide more means of interaction. They are more explorative, collaborative and expressive. In recent years, research in the area of alternative computer environments (especially tangible interfaces) for children has been increasingly popular. But so far work on the evaluation of tangible technologies has been rather scarce [5].

Originally this project was inspired by the use of physical items in 'museum loan boxes'. These boxes are supplied by museums on request to schools, and contain replica objects which can be explored in the classroom. To investigate how the artefacts in the loan boxes could be made more interactive, we aim to design and evaluate tangible technology based on the loan box items by using RFID technology.

We are asking the following research questions:

- Will TUIs encourage children's involvement in the learning activities and therefore motivate them to learn?
- Will TUIs provide easier access to learning materials than traditional computer interfaces?

2. MAIN CONTRIBUTIONS

The research will contribute to paradigms such as the design and evaluation of 'disappearing computer' and 'tangible computing'. In-depth comparison between evaluation methods and discovering of new method(s) could be beneficial to designing TUIs for children. The results will be used for developing evaluation guidelines.

3. DESCRIPTION OF THE METHODOLOGY ADOPTED

We work with teachers and children (key stage 2) in real world contexts: in the classroom, in the home and in informal learning contexts such as museums and exploratoria.

According to the literature review, many methods have been used for evaluating products for children [1, 2, 3, 4]. In my study, initially the evaluation was conducted using Think-Aloud (TA), Peer Tutoring (PT), Drawing Intervention (DI) and Questionnaires.

4. CURRENT STATUS

A review of Tangibles and human-computer interaction on the design of tangible technologies for learning has been done. We also carried out initial observational studies on the use of physical and/or digital manipulatives in primary school classrooms.

After some school classroom observations and design workshops with children, the first prototype with RFID technology was built for the initial evaluation. It is a puzzle game on the life of the Romans. Children are given a number of different Roman replicas to investigate. Some of the items may have been used by the Romans in their dining room. When children select an item and place it into the miniature Roman Room, information on the item will be given in the form of visual and audio feedback.

We aimed to evaluate mainly the usability of the prototype, and also fun and educational design. The evaluations were carried out in both school and labs. We selected some user-based evaluation methods and carried out formative evaluation studies with this early TUI prototype and drawn some initial conclusions on the evaluation methods we used.

5. INTERIM CONCLUSIONS

Each evaluation session gave some useful results about usability issues, mainly regarding some minor flaws in the programming of the interface, and the amount of information delivered to the children, in what form and in what pace.

More interesting findings were about the evaluations and the evaluation methods we used:

- The location plays a large part in how children behave; children felt more at ease and focused when in the school ICT room.
- The disposition of the space (i.e. reading screen, listening to audio, looking at the objects) to avoid the observer sitting in the way [3].
- For Think-Aloud method, the selection of children is important; not all the children were naturally talkative when they used the product.
- For Peer Tutoring, shy children who didn't talk a lot during the Thinking Aloud showed enthusiasm and engagement in teaching their friends how to use the product and helping them to carry out the same tasks again.
- For Drawing Intervention, a number of children placed a chosen artifact in front of them to do the drawing. In terms of letting participant express themselves, DI was the most successful method used.

6. FURTHER WORK

We are still in search for the best evaluation methods for children's tangible technology, the optimal environment setting for the evaluations, and best way of conducting the evaluations.

The design of the prototype system will be refined, more evaluation will be done with children and results will be compared and documented for further use. We will also try to validate the new evaluation methods we propose.

REFERENCES

[1] Donker, A., and Markopoulos, P., A comparison of think-aloud, questionnaires and interviews for testing usability with children, In: Faulkner, X., Finlay, J. & Detienne, F. (Eds.) "People and Computers XVI - Proceedings HCI 2002", Springer, p305-316.
[2] Hanna, L., Risden, K. and Alexander, K. J., (1997) *Guidelines for usability testing with children*, Interactions, (5) 9-14.
[3] Höysniemi, J., Hämäläinen, P., and Turkki, L. (2003). *Using peer tutoring in evaluating the usability of a physically interactive computer game with children.* Interacting with Computers, 15(2), 205-225.
[4] MacFarlane, S., Read, J., Höysniemi, J. and Markopoulos, P. (2004). *Evaluating Interactive Products for and with Children*. Tutorial presented at CHI2004, Vienna, Austria.
[5] O'Malley, C. and D. S. Fraser (2004). *Literature Review in Learning with Tangible Technologies*, NESTA Futurelab. http://www.futurelab.org.uk/download/pdfs/research/lit_reviews/futurelab_review_12.pdf
[6] Xu, D., Mazzone, E. and MacFarlane, S. (2005). Informant Design with Children - Designing Children's Tangible Technology, 1st International Workshop "Re-Thinking Technology in Museums", Limerick, Ireland http://www.uclan.ac.uk/facs/destech/compute/staff/read/Publish/ChiCi/references/informant_design_with_children.pdf

Affective Posture Recognition: Human Factors and Modelling

Andrea Kleinsmith
Department of Information Systems
University of Aizu, Aizu-Wakamatsu, Japan
andi@andisplanet.com

Nadia Bianchi-Berthouze
Department of Computer Science
University College London, London, UK
n.berthouze@ucl.ac.uk

Through technology, it has become increasingly important to create more natural and effective interactive interfaces that are capable of communicating with the user through affective channels to improve user experience. The role of body posture in affect recognition, and the importance of emotion in the development and support of intelligent and social behavior have been widely accepted. Our focus is on creating an affective posture recognition system that incrementally learns to recognize users' affective behaviors, such as posture. This information may then be used to inform another application. Thus far, my research has progressed in three main directions. Towards understanding to what extent culture and gender may be important for recognizing affective states, what aspects of affect should be modelled and how, and the identification of a set of low-level, posture features.

Affective posture recognition, adaptive models, affective dimensions, discrete emotion categories, appraisal process theory of emotion

1. INTRODUCTION

Through technology, it has become increasingly important to create more natural and effective interactive interfaces that are capable of communicating with the user through affective channels to improve user experience. These systems also should be able to increase user engagement to make the interaction more effective. While we acknowledge facial expression [6] and voice [2] as important affective channels, research in these fields is quite extensive and continues to be a major focus, whereas the study of whole body postures remains a novel area of research within computer science [4]. Indeed, while there are formal models for classifying affective facial expressions [6], there are no equivalent models for affective posture. Moreover, studies in cognitive neuroscience, psychology, and human-computer interaction have shown the importance of body posture in affect recognition [7] [11], and that the body is used for emotional display more than formerly thought [1] [3].

2. RESEARCH ISSUES

My research goal is to design an affective posture recognition system that adapts to the user, over time, through interaction. We plan to test the system in an educational scenario. In order to create a system that exploits posture in affective communication, several research questions need to be addressed:

1. To what extent do culture and gender affect the recognition of affective posture?
2. What should be modelled and how?
3. How to create recognition models online and how to bootstrap new models from existing models?
4. What should the evaluation metrics be in an affective recognition system?

3. STATUS

Where am I? My progress thus far has been in three main directions. I began by evaluating how culture [9] and gender [8] may affect the personalization of an affective posture recognition system. We created a set of affectively expressive avatars from motion capture data, and conducted a series of online posture evaluation surveys. The use of avatars allowed us to avoid possible confounds by giving us the ability to eliminate various factors, such as face, gender, culture, and age.

Towards the second research question, we have examined the possibility of mapping affective postures onto a set of affective dimensions [10]. Previous approaches to modelling emotion are limited because they categorize expressions into discrete emotion categories, such as angry, happy, sad. A major limitation to using this type of approach in some situations is that the emotional behaviors exhibited may be more subtle and thus, may not fall into distinct and well-separated emotional categories. Instead, the behaviors expressed may encompass a much wider range of feeling and emotion states.

In addition, we used various statistical methods and the results from our surveys to try to determine how and what should be modelled from affective posture. We identified a set of low-level kinematic posture description features to classify posture according to the affective dimensions discussed above [3].

4. INTERIM CONCLUSIONS

The statistically evaluated results from our culture and gender studies found that, for both factors, significant differences do exist. In trying to understand how and what should be modelled, we were able to determine a set of affective dimensions onto which postures could be mapped. Additionally, these dimensions could be grounded into our set of low-level posture features, and they seem to reflect the criteria that human observers exploit when evaluating affective postures.

5. FUTURE RESEARCH

We can not think that one posture equals one emotion. Instead, the emotion is the culmination of several appraisal steps and postures [12]. Considering emotion as an outcome of an ongoing appraisal of a specific event may allow us to address the complex nature of emotion that transcends beyond the set of basic affective states that traditionally has been proposed. In agreement with Scherer, we believe that by using an appraisal process theory of emotion, both the physical and the behavioral significance of different events can be determined as they occur. According to Scherer's appraisal theory, the body (as well as other modalities) changes in response to the different appraisals. How can we use posture to infer the reasons behind the various behaviors and expression changes (represented as a sequence of body postures) as they occur so that we can reconstruct the user's appraisal process?

Obtaining feedback about the state of the user in online models is important. But how can this feedback be effectively obtained? It is not feasible to constantly interrupt the user to obtain feedback from her. Furthermore, it is likely that the user is not always consciously aware of her affective state at each moment in performing a task.

REFERENCES

[1] Argyle, M. (1988) *Bodily Communication - 2nd edition*, Methuen Co. & Ltd.
[2] Bachorowski, J. (1999) Vocal expression and perception of emotion, *Current Directions in Psychological Science*, 8, 53-57.
[3] Bianchi-Berthouze, N. and Kleinsmith, A. (2003) A categorical approach to affective gesture recognition, *Connection Science Special Issue on Epigenetic Robotics – Modeling Cognitive Development in Robotic Systems*, 15, 259-269.
[4] Coulson, M. (2004) Attributing emotion to static body postures: recognition accuracy, confusions, and viewpoint dependence, *Journal of Nonverbal Behavior*, 28, 117-139.
[5] Coulson, M. (in press) Expressing emotion through body movement: A component process approach, To appear in L. Canamero & R. Aylett (eds.) *Animating Expressive Characters for Social Interaction*, John Benjamins Publishing Co., Advances in Consciousness Research Series.
[6] Ekman, P. and Friesen, W. (1978) *Manual for the Facial Action Coding System*, Consulting Psychology Press.
[7] Ghijsen, M., Heylen, D., Nijholt, A., and op den Akker, R. (2005) Facial affect displays during tutoring sessions, *IUI 2005 Workshop on Affective Interactions: The Computer in the Affective Loop*, San Diego, CA.
[8] Kleinsmith, A., Bianchi-Berthouze, N. and Berthouze, L. (2006) An effect of gender in the interpretation of affective cues in avatars, *AVI 2006 Workshop on Gender and Interaction: Real and Virtual Women in a Male World*, Venice, Italy, 23-26, May.
[9] Kleinsmith, A., De Silva, P.R., and Bianchi-Berthouze, N. (in press) Cross-cultural differences in recognizing affect from body posture, *Interacting with Computers*.
[10] Kleinsmith, A., De Silva, P.R., and Bianchi-Berthouze, N. (2005) Grounding affective dimensions into posture features, *Proceedings of First International Conference on Affective Computing and Intelligent Interaction*, Beijing, China, 23-25 October, 263-270.
[11] Meeren, H., van Heijnsbergen, C. and de Gelder, B. (2005) Rapid perceptual integration of facial expression and emotional body language, *PNAS*, 102, 16518-16523.
[12] Scherer, K. (2001) Appraisal considered as a process of multi-level sequential checking, In K. Scherer, A. Schorr & T. Johnstone (eds.), *Appraisal Processes in Emotion: Theory, Methods, Research,* Oxford University Press.

Workshops

(re)Actor: The First International Conference on Digital Live Art

Jennifer G. Sheridan
Computing Department, InfoLab21, South Drive,
Lancaster University, Lancaster LA1 4WA
sheridaj@comp.lancs.ac.uk

Alice Bayliss
School of Performance and Cultural Industries,
University of Leeds,
West Bretton, Near Wakefield, WF4 4LG
a.bayliss@leeds.ac.uk

***Digital Live Art* is the intersection of human-computer interaction (HCI), live art and computing. This conference seeks to bring together practitioners and academics from these diverse backgrounds for a lively debate and interactive event which will explore this emerging field. Our specific context focuses on club cultures as a living context for Digital Live Arts practices. Our expected outcomes are to create a community of Digital Live Artists and to present the first strategies for designing, developing and evaluating Digital Live Art. Such an event provides an opportunity to open up conversations between digital art and live performance and will allow us to explore how it is used to increase our understanding of human-computer interaction in general.**

Digital live art, playful arenas, human-computer interaction, performance art, non-task-based computing, ubiquitous and pervasive environments, club culture

1. WHAT IS DIGITAL LIVE ART?

New technologies enable new types of interaction. While existing human-computer interaction (HCI) focuses on usability, functionality and efficiency, new technologies offer a unique approach to HCI and are different in terms of not only their goal, but also their purpose and evaluation style. Current research in HCI is beginning to look at the interaction that takes place at the intersection of live art and computing [1, 2, 3]. We call this intersection Digital Live Art.

Digital Live Art is the intersection of live art, HCI and computing. *Live Art* is a term that is often understood in relation to its more popular parent Performance Art which emerged as an “unconventional” art form after Allan Kaprow coined the term *happenings* in the 1960s. It focuses on presence or liveness: the live artist, her body and her bodily actions rather than on material objects [4], as well as the relationship between the artist and audience. In Live Art, the performer *is* the artist [5] and often the artwork itself. Current performance research suggests that Live Art is a complex weave of the live, immersive and interactive [6] that reconnects art and life [4] not simply through the everyday, but from a magnification and reconfiguring of the everyday [7].

There is a seeming contradiction that exists between that which is digital and that which is live. In many cases, Digital Art is not 'performed' in the traditional sense of the word and is often communicated, made or received via the machine rather than through living human-human encounters. On the other hand, more and more artists are utilising digital media as a way of conceiving, producing and mediating performed art works. It is our view that instigating the possible conversations that might exists between the digital and live modes is critical.

Digital Live Art is a hybrid art form which focuses on presence and presupposes the digital as a way of making live engagements. This liveness is mediated by technical dependency so that the performance becomes a hyper-real exploration of the mundane - an orchestrated, temporal event occurring for any length of time and in any place using technological means. This technical relationship is negotiated within a malleable yet identifiable performance frame [8, 9] and context.

Our particular interest is in exploring the relationship that develops between performers, participants and observers and how Digital Live Art moves people to performative interaction and communal engagement. Whilst the types of ubiquitous and pervasive environments being explored in Digital Live Art are multiple and varied, our particular interest is in the increasingly popular but extremely challenging ubiquitous environment called a *playful arena [2].*

2. CLUB CULTURE AS LIVING CONTEXT

Playful arenas such as nightclubs and festivals, are spaces that already bristle with technology - computers, decks, mixers, projectors, lights, webcams and so on. The technology is utilized to fulfil a basic commitment to providing physical, sensory and human communication, interaction and experience for groups of people pursuing pleasure. As such it is a particularly rich environment for positioning Digital Live Art that seeks to explore the shifting binaries of performer/audience, human/machine, creator/collaborator.

The bringing together of the seemingly disparate strands of interactive performance, club culture and sensor technology reflects our own working practice that has developed over the past four years. Our collaborative research comes variously from the world of theatre and performance and from the world of computing. Whilst these very distinctive fields have particular methodologies, fields of reference and specific vocabularies, we link them philosophically with our desire to explore the nature of human interaction and our belief that a playful arena is a fertile ground for experimentation, innovation and creativity. We seek to investigate how Digital Live Art can lead individuals and groups to engage with or participate in the construction of their world and perhaps begin to take ownership of it. As Schieffelin [10] says:

> Without living human bodily expressivity, conversation and social presence, there would be no culture and no society…performativity is not only endemic to human being-in-the-world but fundamental to the process of constructing a human reality.

It is a widely held view that the club space can exist as a ‘playground of culture’ [11] where the distinctions between performer and spectator, between those who act or take action and those who do not, have already become blurred. Drawing on experiences and influences from the DIY party movement of the 1990s, the underground club supports a ‘kind of terrain, a shifting dance environment without borders or destination’ [12, pp 173] where the gaze turns away from the brightly lit, bounded stage world of the professional performer and back towards the crowd. As this clubber testifies, performative energy emanates from the floor:

> The very dynamic of the rave itself felt so liberating - democratic rather than hierarchical. The dancers’ focus was not on the stage, but on each other. The hegemony of the performer was usurped, the energy was coming from the participants themselves. [13, pp. 104]

It is in support of performative democracy that we seek to position our Digital Live Art and share its possibilities. The aim is for it to sit alongside and complement the aesthetic of the club already established by the clubbers themselves, rather than replace or usurp it. We are invited into the club to play as artists and we invite people to interact with the artwork so that they may experiment and create for themselves and others. As Graham says of interactive art and its underlying principle of facilitating the participants’ experience:

> The skills needed by artists if they are to truly loosen control over the audience, but still share their pleasure, are perhaps less like the traditional art skills, and more like the social interaction skills of ‘throwing a good party’, or of enabling community art. [14, pp.171]

Using Digital Live Art in playful arenas introduces new concepts, methodologies and applications for live performance that may have wider implications for tangible and wearable computing, pervasive environments, ambient intelligence and collaborative creativity beyond the club space. It is our intention to bring together people who are interested in mapping this exciting cross-disciplinary field.

REFERENCES

[1] Sheridan, J. G. (2006) Digital Live Art [PhD Thesis]. Lancaster: Computing Department, Lancaster University.
[2] Bayliss, A., Lock, S., Sheridan, J.G. (2004) Augmenting Expectation in Playful Arena Performances with Ubiquitous Intimate Technologies. Proceedings of PixelRaiders 2, Sheffield, 6-8 April (CD-ROM).
[3] Sheridan, J.G., Dix, A., Bayliss, A., Lock, S. (2004) Understanding Interaction in Ubiquitous Guerrilla Performances in Playful Arenas. In S. Fincher, P. Markopolous, D. Moore, & R. Ruddle (eds), People and Computers XVIII-Design for Life: Proc. of HCI 2004. Springer-Verlag.
[4] Schimmel, P. (Ed.) (1998) Out of Actions: Between Performance and the Object 1949-1979. Thames and Hudson, London.
[5] Goldberg, R. (2001) Performance Art: From Futurism to Present. Thames and Hudson, London.
[6] Heathfield, A. (Ed.) (2004) Live: Art and Performance. Tate Publishing, London.
[7] Schimmel, P. (Ed.) (2003) art, lies and videotape: exposing performance. Tate Publishing, London.
[8] Bateson, G. A (1955) Theory of Play and Fantasy. Psychiatric Research Reports., 2, 39-51.

[9] Goffman, E. (1974) Frame Analysis: An essay on the organization of experience. Northeastern University Press, New York.
[10] Schieffelin, E. L. (1998) Problematizing performance. In Hughes-Freeland, F. (ed), Ritual, Performance, Media. Routledge, London.
[11] Buckland, F. (2002) Impossible Dance: Club Culture and Queer World Making. Wesleyan University Press. Conneticut,
[12] Reynolds, S. (1990) Blissed Out: The Raptures of Rock. Serpent's Tail, London.
[13] Collin, M. (1997) Altered State: The Story of Ecstacy Culture and Acid House. Serpent's Tail, London.
[14] Graham, B. (1996) Playing with Yourself: Pleasure and Interactive Art. In Dovey, J. (ed) Fractal Dreams: New Media in Social Context, Lawrence and Wishart, London.

Combining Visualisation and Interaction to Facilitate Scientific Exploration and Discovery

Elena Zudilova-Seinstra
Scientific Visualisation and Virtual Reality Group
Section Computational Science, University of Amsterdam
403 Kruislaan, 1098 SJ Amsterdam, The Netherlands
elenaz@science.uva.nl

Tony Adriaansen
ICT Centre, CSIRO,
P.O. Box 76, Epping, NSW 2121
Australia
tony.adriaansen@ csiro.au

This one-day workshop is aimed to facilitate the information exchange on theoretical, generic and applied aspects of HCI within the broad scope of scientific and interactive visualisation. We will offer an interdisciplinary forum of discussion for practitioners interested in designing interactive visualisation systems and academic researchers from HCI, computer science, artificial intelligence, software engineering, sociology and psychology. The ultimate goal of this workshop is to bridge the communication gap that exists today between HCI experts and those developing visualisation algorithms and systems.

Intuitive exploration environments, interactive and collaborative visualisation, design guidelines, usability

1. INTRODUCTION

With the explosion of IT the amount of data at one's disposal is enormous. The data explosion has led to very large detailed datasets and the amount of details in these datasets continues growing at explosive rates. Nowadays the challenge is to harness the power of knowledge hidden in the collage of scientific data.

Visualisation technologies empower users to perceive important patterns in a large amount of data, identify areas that need further scrutiny and make sophisticated decisions [7]. But looking at information is only a start. Users also need to manipulate, explore and share the data with other people. Unfortunately, for many application domains it is even unclear what the features of interest are and how to define them in such a way that they can be detected. As a result the need for direct user interactions is becoming crucial. To increase the users' ability to explore the data and better understand the results of scientific experiments based on extensive calculations new scientific methods need to be developed and applied. The importance of human interaction and perception cannot be ignored.

2. INTUITIVE DATA EXPLORATION

The intuitive exploration of scientific data can only be possible when advanced and sometimes even multi-modal interaction and visualisation technologies are integrated. When combined properly the ability to visualise and interact with the data can aid analysis and understanding in many areas, such as: scientific experiments, manufacturing process control, financial data analysis, etc.

With the rapid development of HCI our work and learning has become more efficient and fascinating. The maturing virtual reality (VR) techniques [1, 3], together with the emerging haptic interfaces and multimedia networking technologies provide ground for new enabling tools addressing computer-supported activities. They open the way to new forms of collaborative work and new domains of multi-participant systems. Today's VR-based collaborative visualisation and exploration environments are becoming the established ways of facilitating human interaction with large amounts of information. They allow opinions to be shared, removing the data bottleneck of individual analysis and reducing the time to discovery.

Multi-modal interaction systems have become very popular as well. Based on the combination of different interaction techniques (i.e., direct manipulation, speech recognition, haptics, real time video and audio, etc.), multi-modal systems [6] aim to provide efficient, convenient and natural interaction and communication between computer systems and users in a seamless way and will ultimately enable people to interact more fully within an exploration environment using everyday skills.

Another important requisite in the interaction design is aimed at emphasizing ‘human-to-human’ properties, so called social user interfaces [5]. This kind of design considers human emotions and personality e.g., face-to-face communication between users, embodied agents, etc. Embodied agents usually interact with users or each other via multi-modal communicative acts, which can be non-verbal (virtual embodied agents) or verbal (conversational embodied agents). They permit building a kind of relationship with an interactive environment as well as with other users to assist in the exploration process.

3. INTERACTIVE VISUALISATION

Even though advanced interaction technologies are applied more and more often to the field of scientific and information visualisation, the integration of visualisation and interaction methods is still an open question. The relatively new concept of interactive visualisation introduced recently aims to address this research concern [4]. The purpose of interactive visualisation is not only to provide users with a possibility to view the data and modify representation parameters but also to permit them using interaction abilities for the interrogation and navigation through datasets and communication these insights with others.

Unfortunately, today’s interactive visualisation still remains focused mainly on high-performance algorithms, efficient feature extraction techniques and human perception parameters [2]. As for HCI and usability aspects, they have not yet been sufficiently addressed. There exists a major communication gap between HCI experts and those developing visualisation algorithms and systems. The former often consider visualisation methods and techniques as tools oriented to computer scientists with very little relevance to HCI. Vice versa, to those working on scientific visualisation HCI concepts still often remain an after-thought rather than an essential component that affects the system’s quality. In fact, new visualisation techniques are rarely compared with previous results and their effectiveness is seldom quantified by user studies. Meanwhile, interactive visualisation systems are supposed to be easy to use and not require sophisticated skills by users, as many are not computer experts.

4. WORKSHOP OBJECTIVES

The aim of this workshop is to promote the importance of HCI and usability aspects in visualising and exploring complex datasets and sharing obtained knowledge with other people. It will serve as an international forum for the information exchange on theoretical, generic and applied aspects of HCI within the broad scope of scientific and interactive visualisation. We expect to attract researchers and practitioners from a wide range of disciplines, including HCI, computer science, artificial intelligence, software engineering, social science and psychology, to discuss how the combination of modern interaction and visualisation technologies may facilitate the information analysis cycle from the user’s perspective.

The main topics to be addressed by the workshop are as follows. How can interactive visualisation methods and tools be augmented to address both concerns of scientific computing and HCI? What are the criteria for choosing between advanced projection equipment and input devices? How can visualisation and multi-modal interaction techniques be integrated to ensure a good user experience? Which research questions need to be considered when aiming to achieve efficient HCI and ‘human-to-human’ interaction between scientists while exploring complex data spaces? In what way users and the environment of use should be modelled when designing interactive visualisation systems? How can we define effective abstractions for the visualisation and user interaction processes? What is the impact of task- or application field-orientation on interactive visualisation? To what extent are usability problems independent of the context of use and need to be taken into account when designing interactive visualisation environments?

The workshop submission will be accepted from the academic, industrial and commercial institutes. We plan to invite both position papers and extended research abstracts that will be reviewed by Program Committee members and selected based on originality, contribution to the goals of the workshop and the shared interest of participants. The workshop will be organised in individual brief presentations, problem-oriented group activities, group presentations and a final round-table discussion of results. The workshop will also offer the possibility to discuss existing interactive visualisation tools and systems, including collaborative ones. The developers will be able to show their products (or research prototypes) to potential users, while users can pose their questions and requests.

REFERENCES

[1] van Dam, A., Laidlaw, D.H., Simpson, R.M. (2002) Experiments in Immersive Virtual Reality for Scientific Visualization, IEEE Computer and Graphics, **26** (**4**), 535-555
[2] Jonhson, C. (2004) Top Scientific Visualization Research Problems, IEEE Computer Graphics and Applications, **4**, 13-17.

[3] Shneiderman B. (2003) Why Not Make Interface Better than 3D Reality? IEEE Computer Graphics and Applications, **23(6)**, 12-15
[4] Nielson G.M., Hamann B. (1990) Techniques for the Interactive Visualization, Proceedings of IEEE Visualization Conference 90, San Francisco, California, 23-26 October, 45-49, IEEE Computer Society Press
[5] Nijholt, A. (2003) Disappearing computers, social actors and embodied agents. Proceedings of 2003 International Conference on CYBERWORLDS, Singapore, 3-5 December, 128-134, IEEE Computer Society Press
[6] Oviatt S., Darrel T. and Flickner M. (2004) Multimodal Interfaces that Flex, Adapt, and Persist, Comm. ACM, January. 2004, 30-33
[7] Spence R. (2000) Information Visualisation, ACM Press

Designing with Elderly for Elderly

Corina Sas, Peter Bagnall
Computing Department InfoLab21, Lancaster University,
South Drive, LA1 4WA UK
c.sas@lancaster.ac.uk, pete@surfaceeffect.com

Alina Coman
Sociology Department, Transylvania University
Bd. Eroilor, nr. 29, Brasov 500036, Romania
A25coman@yahoo.com

This one-day workshop will offer an interdisciplinary forum of discussions for practitioners and academic researchers interested in designing for elderly users. The two important issues that will be explored relate to the distinct set of needs, expectations and values that characterises tomorrow's generation of elderly, the retired baby boomers. These characteristics challenge our preconceptions of older users and require a new approach to design. The second issue pertains to design methodology and seeks to explore innovative design methods that can be used to successfully involve elderly users in the design process.

Elderly users, global aging, design methodology, participatory design

1. INTRODUCTION

Designing for older population is a research topic that has started receiving increased attention within the HCI community. The rationale for this is twofold. On the one hand, the greater life expectancy leads to an increasingly larger segment of older population. On the other hand there is the phenomenon of global aging with the young population decreasing while the older one is steadily increasing. This trend is further accelerated by the baby boomer generation which is starting to retire. Statistics suggest that by 2025, more than a third of the UK's population will be over 55 [1].

Longevity is not the only trend, it couples with the so called compression of morbidity, which refers to the fact that people tend to stay healthier until an old age. These features redefined the concept of elderly and this new type of elderly is the one that HCI community should place emphasis on:

> "UK baby boomers have the wealth, health and spare time to 'live life again'. At the moment, many UK boomers are beginning to enjoy a windfall; the combination of wealth, good health and more and more spare time gives them a new phase of life in middle age that is set to continue into their older years. Increasingly free from the pressures of overwork and childrearing baby boomers have the chance to 'live life again' and satisfy their desire for personal fulfilment" [2].

This socio-demographic scaffold challenges the traditional perception of older population. Given that physical and cognitive abilities decline with age [3], this traditional perception [4] usually sees elderly users as less-abled [4], dependent [5] or isolated [7]. This traditional approach to design for elderly lies mostly under the remit of assistive design.

There appears to be a gap between the tomorrow's generation of elderly users, e.g. retired baby boomers, and the current understanding of elderly users in nowadays HCI design. This workshop aims to address this gap and to challenge the preconceptions about elderly users which may not hold valid for tomorrow's generation of elderly users. A powerful way of addressing this gap is through design methodologies developed for working for and with elderly users [9]. Such design methods will focus on capturing the emotional requirements of elderly people and actively involving them in the design process. Requirements elicitation is a fundamental stage in the process of designing interactive systems. The emphasis on designing for elderly has received increasing attention in the last years, e.g. CHI 2005 student competition focused on designing for supporting the well-being of seniors above the age of 65. Within this area, the emphasis was mainly on designing for elderly, with relative less interest on including elderly people in the design process. Given the individual differences of this group of age, the classic participatory design method may need to be adjusted when the target groups are elderly. In addition, designing technology for this age group has primarily focused on assistive technology, to address the physical or cognitive needs of elderly people [10, 11], whereas the emphasis on their emotional needs [8,12,13] received less attention.

In this light, this workshop aims to address a gap in the existing research field: developing new design methodologies, not only for elderly but with elderly, for capturing and addressing primarily their emotional needs.

The workshop will serve as a forum for exchanging ideas and aims to address the following questions:

- Identifying requirements of tomorrow's elderly users
- Identifying some shared characteristics among elderly users that can successfully inform design
- Identifying preconceptions regarding elderly users
- Identifying design guidelines for technology for elderly users
- Designing games for tomorrow's elderly users
- Design methodology for working with elderly users
- Identifying the specificity of participatory design when applied to elderly users

2. WORKSHOP PROCEDURE

Format: One-day workshop

Workshop Agenda:

9:00-10:30	**Session 1**	
	9:00-9.15	Introduction, overview of workshop agenda
	9:15-10:30	Participants' brief presentations
10:30-11.00	**Break**	
11:00-12:30	**Session 2**	
	11:00-12:00	General discussion focused on the opportunities and challenges pertaining to the design for tomorrow's elderly
	12:00-12:30	Summarising the relevant problems when designing for elderly. Using these findings and those previously identified by the organisers for preparing the afternoon learning group work.
12:30-14:00	**Lunch**	
14:00-15:45	**Session 3**	
	14:00-15:00	Problem-based learning: participants will work on small groups for finding ways to address the identified problems
	15:00-15:45	Group presentations and discussion
15:45-16:15	**Break**	
16:15-17:30	**Session 4**	
	16:15-17:00	Reflecting on workshop activities and conclusions with the purpose of sketching the poster
	17:00-17:30	*Wrap up*

REFERENCES

1. Jeavans, C. (2004) BBC News: Welcome to the ageing future. http://news.bbc.co.uk/1/hi/uk/4012797.stm
2. Harkin, J. and Huber, J. (2004) Eternal Youths: How the Baby Boomers are Having Their Time Again. Demos, London.
3. Salthouse, T.A. (1991) Theoretical perspectives on cognitive aging. N.J.: Lawrence Erlbaum Associates, Hillsdale.
4. Östlund, B. (2003) Social science research on technology and the elderly – Does it exist? The Institute for the Study of Ageing and Later Life. University of Linköping, Norrköping Campus. www.certec.lth.se/britt.ostlund/SocialScience.pdf
5. Marcus, A. (2003). Fast forward: Universal, ubiquitous, user-interface design for the disabled and elderly. Interactions, 10(2), 23-27.
6. Hirsch, T., Forlizzi, J., Hyder, E., Goetz, J., Kurtz, C. and Stroback, J. (2000) The ELDer project: social, emotional, and environmental factors in the design of eldercare technologies. Proceedings of Universal Usability. ACM Press.
7. Hollywood, E., O'Brien, G., Lennon, S. (2005) Student design competition: SIGCHI project: user centered design of a program alleviating loneliness (PAL). CHI '05 Extended abstracts on Human factors in computing systems
8. Santana, P.C., Rodríguez, M.D., González, V.M., Castro, L.A. and Andrade, A.G. (2005) Student design competition: Supporting emotional ties among Mexican elders and their families living abroad. CHI '05 Extended abstracts on Human factors in computing systems

9. Fisk, A. D., Rogers, W. A., Czaja, S. J., Charness, N. and Sharit, J. (2004). Designing for older adults: Principles and creative human factors approaches. CRC Press, Boca Raton.
10. Jorge, J.A. (2001). Adaptive tools for the elderly: new devices to cope with age-induced cognitive disabilities. Proceedings of the 2001 EC/NSF Workshop on universal accessibility of ubiquitous computing: providing for the elderly. 66-70
11. Coleman, K. (2001). Electromyography based human-computer-interface to induce movement in elderly persons with movement impairments. Proceedings of the 2001 EC/NSF Workshop on universal accessibility of ubiquitous computing: providing for the elderly. 75 - 79
12. Rowan, J. and Mynatt, E.D. (2005). Digital Family Portrait Field Trial: Support for Aging in Place. CHI 2005: 521-530.
13. Kenyani, P., Hsieh, G., Mutlu, B., Easterday, M., Forlizzi, J. (2005), DanceAlong: Supporting Positive Social Exchange and Exercise for the Elderly Through Dance. Extended Abstracts of CHI 2005.

HCIEd.2006-2 Workshop: Developing the "Yellow Book" of HCI Referent Problems

William Wong[1], Janet C. Read[2] and Paul Englefield[3]
[1]Interaction Design Centre, School of Computing Science, Middlesex University, London, UK.
[2]Child Computer Interaction Group, Department of Computing, University of Central Lancashire, Preston, UK.
[3]Usability Competency Centre, IBM United Kingdom Ltd, Warwick, UK.
w.wong@mdx.ac.uk; jcread@uclan.ac.uk; paul_englefield@uk.ibm.com

This workshop is part of the series of BCS HCI Group's workshops on HCI Education and Practice. The purpose of this year's workshop is to develop a set of referent problems for use in HCI education – both in academia and industry. These referent problems will not replace a curriculum, but it is intended that the problems will be presented in a way that would allow educators to use them directly in their teaching to supplement their curriculum. This workshop follows from the "Inventivity" workshop held in March 2006 in Limerick, and jointly organised by the BCS, the ICS, IFIP and the CONVIVIO network.

HCI education, problem-based learning, design problems

1. INTRODUCTION

The one-day workshop is intended to provide a forum for HCI Educators in both industry and academia to extend their understanding of curriculum development and delivery issues in teaching HCI. This is part of a two-stream series of workshops held twice a year, once in March, and the second at the annual HCI conference in September. The March workshops are designed as agenda setting workshops, while the September workshops will be geared towards developing the resources to support the HCI education agenda.

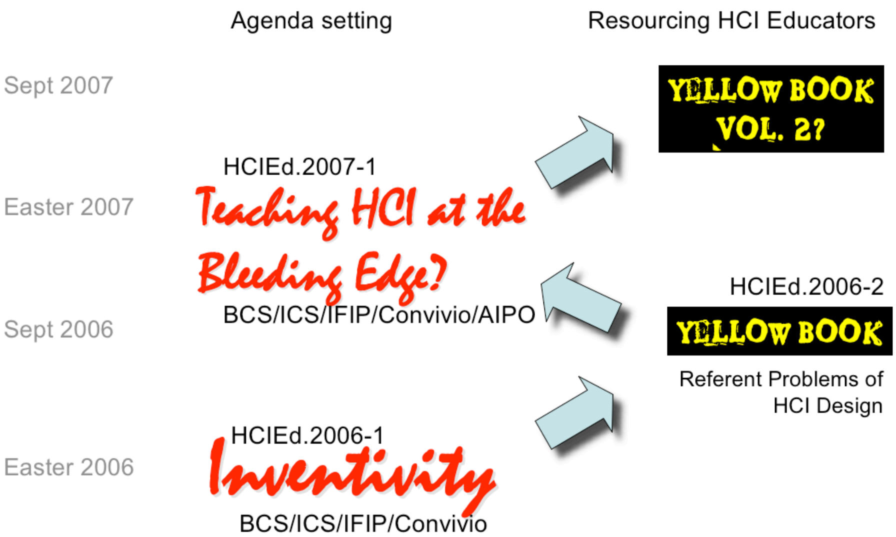


Figure 1. Positioning of the workshops.

In this September, this round of the Workshop series will be looking at the problems that occur in connecting and communicating with people, and the use of that focus to develop a corps of problems that can guide curriculum development in HCI. We hope to use that as a focus to develop a "Yellow Book" of Referent Problems in HCI.

2. GOAL OF THE WORKSHOP

The goal of the workshop is to develop the "Yellow Book" of HCI referent problems and issues. Participants at the workshop will work towards identifying some "standard" problems where educators and trainers can use to teach different aspects of HCI analysis, evaluation, design and invention. In particular, we will make a first attempt at (i) classifying interaction / usability problems that would guide the kinds of problem sets to develop, and (ii) expressing them in different problem cases that could be used for design exercises in the classroom.

3. CASE STUDIES AND PROBLEMS

Just as there are standard 'referent' problems in programming that illustrate key concepts e.g. Tower of Hanoi for sorting algorithms, it would be of value to develop a similar set of referent problems for the different aspects of HCI design. Such problem cases would be developed along the lines of case studies, similar to those from the Harvard Business School, but instead intended for HCI design. Such problem cases would have enough detail for a student to get their teeth into it. These problems would encompass sufficient contextual information to represent some of the complexity that is present in the real-world. Through such cases, we anticipate helping our HCI students learn to cope with complexity.

These case studies or problems are not intended to be a collection of successful case studies or case studies which represent solved problems or implementations of systems from which to lessons learnt have already been extracted and reported, e.g. cases such as the Olympic Messaging System. These cases should present students with opportunities to identify the problem to be addressed and to define it such that a solution can be developed.

These problem cases could also be classified to show classes of interaction and visualisation design problems which students need to work on in order to demonstrate their competence and understanding of how to apply HCI principles, concepts, theories or guidelines in a non-mechanical way. These problems would also embed the socio-organisational issues which often influence design decisions.

4. OUTCOMES FROM THE WORKSHOP

Technology has over the last seven or so years advanced very quickly, developing new areas of application such as mobile computing, communications and e/m-commerce, smart homes, ubiquitous and embedded computing, collaborative systems, higher integration between process control and technology, distance education and e-learning. While the basic HCI curriculum has remained fundamentally intact, its applications into these new and different areas have changed. For an educator to keep up with the changes is quite a challenge. The outcome from this workshop will provide educators with a resource that will help explore HCI students understand the impact of their analysis, design and development, in more real world type of problems than is available today.

There will be a set of 4-6 page proposed HCI problems to be submitted by each participant which will be compiled into a set of workshop proceedings. The target output will be to publish the "Yellow Book" of Referent Problems in HCI Design in the form of a post-conference proceedings as an initial “prototype” of the Yellow Book. It is anticipated that each case may be of between 10-16 pages to provide adequate material for a 2-3 week on-going classroom discussion. Its adoption and use could then be discussed at the following March workshop.

The First International Symposium on Culture, Creativity and Interaction Design

Peter Wright
York University
peter.wright@cs.york.ac.uk

Ann Light
Queen Mary University of London
annl@dcs.qmul.ac.uk

Janet Finlay
Leeds Metropolitan University
J.Finlay@leedsmet.ac.uk

The culture, creativity and interaction design symposium seeks to strengthen the dialogue between the diverse disciplines contributing to culture, creativity and interaction design research and will provide a forum for interdisciplinary analysis and experiences of digital media. It will focus on the role of technology in enhancing culture and creativity. It will seek critical and reflective approaches to the design and analysis of interactive technology. It will also host an exhibition of interactive art.

Culture, creativity, interaction design

1. BACKGROUND

In 2004 the EPSRC funded the *LeonardoNet* network under their call for network proposal in the area of culture creativity and interaction design. The motivation for the network was to develop an interdisciplinary community of researchers to propose a programme of research at the intersection of human-computer interaction, the arts and the humanities. The aim of LeonardoNet is to establish an internationally leading research programme in the rapidly expanding area of Culture, Creativity and Interaction Design. The First International Symposium marks the culmination of this project.

2. MOTIVATION

Advances in interactive computing technology have blurred the line between art and science. Individuals working in multimedia fields must increasingly have an understanding of both domains. This is reflected in company names like Lucas film's "Industrial Light and Magic" and the Disney sub-field of "Imagineering". Ongoing advances in computing technology also offer broadening access to art and new means of appreciating older forms. The digitisation of an art collection not only means that access can be widened if it is presented online but also that the pictures can be viewed in totally new ways. Our appreciation of certain works can also be deepened with the application of interactive technologies.

The developments outlined above have been of interest to researchers from the arts, humanities and sciences. This interest spans theoretical and practical issues and is bi-directional. Recent work from the humanities and arts has constructively critiqued traditional interaction design theory and practice [1,2] but developments in HCI and cognitive science also promise to provide languages and frameworks for exploring the potential of interactivity in contemporary arts and performance, as well as providing new tools for creativity.

Obvious manifestations of this new radical-interdisciplinarity is the concern in HCI with experience-centred design, and the growing awareness of the need to understand the relationships between the aesthetics and functionality of digital devices [3,4]. Theories and methods from the humanities enrich understanding of concepts such as audience reception and experience, which can be applied in interactive systems design. In addition, there is a growing awareness of the digital art movement and analysis of interactive systems as digital media, which opens up a space for new ideas about how to frame interaction design and the possible applications of interactive technologies.

3. AIMS OF THE SYMPOSIUM

The aim of this symposium is to strengthen the dialogue among the diverse disciplines contributing to culture, creativity and interaction design research and will provide a forum for interdisciplinary analysis and experiences of digital media. It will focus on the role of technology in enhancing culture and creativity. It will seek critical and reflective approaches to the design and analysis of interactive technology.

Topics will include:

- Arts-HCI;
- New media and genres;
- Technology and experience;
- Enhancing creativity;
- Performing arts; I
- Identity politics;
- Critical theory;
- Experience-centred design,
- Place space and interactivity;
- Experiences of working across arts-science boundaries.
- Interactive exhibits

4. ORGANISATION OF THE DAY

The symposium will include two international keynote speakers and a series of shorter presentations. There will also be an exhibition of interactive art at the symposium.

The selection of short presentations will be based on extended abstracts. Abstracts will be reviewed by a programme committee and accepted abstracts will be published in the symposium proceedings.

Our previous experiences of organising workshops at CHI, Interact and DAC lead us to stress the importance of discussion time and the programme will be designed to maximise this. In addition since our aim is to strengthen dialogue across disciplines and to build a community of researchers, non-presenting participants will be encouraged to attend.

REFERENCES.

[1] Bolter, J.D., Gromala, D. (2003) Windows and Mirrors: Interaction design, digital art and the myth of transparency. MIT Press.

[2] Coyne, R. (1999). Technoromanticism: Digital Narrative, Holism, and the Romance of the Real. Cambridge, Mass.: MIT Press.

[3] Norman, D. (2004). Emotional design: Why we love or hate (everyday) things. Basic Books.

[4] McCarthy, J. and Wright, P.C. (2004) Technology as Experience.

Engaging with Emotions - the Role of Emotion in HCI

Christian Peter
Fraunhofer Institute for Computer Graphics
Joachim Jungius Str. 11
18059 Rostock, Germany
cpeter@igd-r.fraunhofer.de

Elizabeth Crane
University of Michigan
401 Washtenaw Ave.
Ann Arbor, MI 48108, USA
bcrane@umich.edu

Lesley Axelrod
School of Information Systems Computing and Mathematics
Brunel University,Uxbridge, Middlesex, UB8 3PH, UK
lesleyaxelrod@brunel.ac.uk

Russell Beale
Advanced Interaction Grou
School of Computer Scienc
University of Birmingham
Birmingham B15 2TT UK
R.Beale@cs.bham.ac.uk

Emotions are of increasing interest to the HCI community and there is an enthusiastic spirit of adventure amongst researchers exploring this area. Much technical research has been done focussing on development of various proof-of-possibility studies and prototypical emotion components, but work done to date has been fragmented and lacks coherence. The goal of this workshop is to share information among researchers and practitioners working in this field through organized discussion and thematic working groups. This workshop will continue bringing people together, forming a community, supporting networking and collaboration to advocate coordinated work in this fascinating field of research.

Emotion, HCI, affective computing, emotion detection, affective applications

1. INTRODUCTION

The topic of emotion in Human-Computer Interaction is of increasing interest to the HCI community. Since Rosalind Picard's fundamental publications on affective computing [1-4], research in this field has gained significant momentum. More and more HCI researchers and application designers are beginning to recognize the importance of emotion in everyday interactions with people and computers alike. Several applications already try to appreciate their users' emotions. Be it by allowing the users to choose a pleasing look for their applications [5], designing Web pages targeted at specific user groups [6], or providing different ways to give feedback or offer help [7], emotions are finally recognised as a key attribute of users which must not be neglected.

Emotion research is largely grounded in psychology yet spans across numerous other disciplines. The challenge of such an interdisciplinary research area is developing a common vocabulary and research framework that a fast developing discipline requires. Emotion related activities in HCI are so far confined to two main areas: testing whether it is possible to "measure" emotion, and if and how emotions could be included in existing applications or artefacts. What is increasingly needed for advanced and serious work in this field is to place it on a rigorous footing, including developing theoretical fundamentals of HCI-related emotion research, understanding emotions' function in HCI, ethical and legal issues, and the practical implications and consequences for the HCI community.

The first workshop on emotion in HCI held in Edinburgh last year [8] brought an interdisciplinary group of practitioners and researchers together for a lively exchange of ideas, discussion of common problems, and identification of domains to explore. The workshop participants identified four key themes to address for continued growth in this domain. The identified themes are:

- theoretical fundamentals such as emotion models, emotion representation, and use of results from psychology;
- emotion detection affairs such as sensing technologies, data analysis, challenges and special requirements of HCI;
- affective applications, particularly why, what, and how to implement them, challenges and demands; and
- ethical and legal issues.

2. WORKSHOP TOPICS

The aim of this year's workshop is to bring together and support the growing community in the field of emotion-related HCI research. The emphasis will be on discussion and joint work on selected topics.

Participants will engage in developing further the themes from the first workshop in as wide an application spectrum as possible, such as internet applications, ambient intelligence, office work, control rooms, mobile computing, virtual reality, presence, or home applications.
Topics addressed by the workshop are:

- How do applications currently make use of emotions?
- What makes applications that support affective interactions successful?
- How do we know if affective interactions are successful, and how can we measure this success?
- What value might affective applications, affective systems, and affective interaction have?
- What technology is currently available for sensing affective states?
- How reliable is sensing technology?
- Are there reliable and replicable processes to include emotion in HCI design projects?
- What opportunities and risks are there in designing affective applications?

It is expected that at the end of the workshop a deeper understanding of the impact of emotion in the wide field of human-computer interaction will have been developed, chances and challenges will have been identified, possible consequences for the HCI community will have been discussed, and first steps to building the theoretical foundations for serious HCI-related emotion research will have been defined. As direct output of the workshop it is intended to publish a special issue of a journal on this topic.

3. WORKSHOP PROCEDURE

The anticipated outline for the one day workshop is as follows:

Introduction: Each participant will give a short introduction to his/her background and position paper
Demos: A slot for demonstrations of working prototypes of affective applications, sensors, data analysis tools and other related work.
Discussion: A discussion of issues raised in the introductory part will be held. Based on the thoughts and expectations canvassed before the workshop and on questions raised during the introduction, the discussion will be led with the goal of identifying themes and goals for the subsequent working groups.
Thematic working groups: Groups will form to work on the topics identified in the discussion. It is planned to work on exemplary case studies to streamline the discussions.
Create outputs: By debating the findings of the working groups, we will aim to develop tangible deliverables that are consensually agreed. These might be:

- Joint publications
- Collaborations
- Networking activities such as setting up online discussion groups/mailing lists, web site, BSCW
- Grant proposals

Defining next steps: At the end of the workshop, the participants will agree on consequent common activities, such as papers to write, projects to develop, the next workshop, or proposing a special issue of a journal.

5. WHO SHOULD ATTEND?

This workshop is intended for both academics and practitioners. Interested participants should submit an extended abstract (about 800 words). Accepted contributions will have the chance to be extended to short papers (4 pages). Case study papers describing current applications or prototypes are strongly encouraged. As a way of bringing the domain to life, presentations of products or prototypes that participants have been involved in are highly encouraged as well. Papers will be reviewed by the workshop's committee members. Workshop updates and accepted papers will be made available on the workshop website.

For more information on the workshop, please see **www.emotion-in-hci.net**

6. WORKSHOP COMMITTEE

Christian Peter, Fraunhofer IGD Rostock, Germany
Lesley Axelrod, Brunel University, UK
Elizabeth Crane, University of Michigan, USA
Dr. Russell Beale, University of Birmingham, UK
Prof. Ruth Aylett, Heriot-Watt University Edinburgh, UK
Prof. Barbara Fredrickson, University of North Carolina, USA
Prof. Kristina Höök, Swedish Institute of Computer Science, Sweden
Prof. Bodo Urban, Rostock University, Rostock, Germany
Prof. John Waterworth, Umeå University, Sweden
Dr. Nicola Millard, British Telecom plc, UK
Dr. Karina Oertel, Fraunhofer IGD Rostock, Germany
Dr. Jörg Voskamp, Fraunhofer IGD Rostock, Germany
Dr. Robert Ward, University of Huddersfield, UK
Steffen Mader, Fraunhofer IGD Rostock, Germany

REFERENCES

[1] Picard, R.W. (1997). Affective Computing. M.I.T. Press, Cambridge, MA.
[2] Picard, R. W, Healey, J. (1997). Affective Wearables, Personal Technologies Vol 1, No. 4 , 231-240.
[3] Picard, R.W.(1999). Affective Computing for HCI. Proc. of the 8th International Conference on Human-Computer Interaction: Ergonomics and User Interfaces-Volume I. Lawrence Erlbaum Associates, Inc.
[4] Picard R.W., Vyzas E., Healey J. (2001). Toward Machine Emotional Intelligence - Analysis of Affective Physiological State. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 23 No. 10.
[5] Brinkman W.-P., Fine N. (2005). Personalising emotional skin designs. Workshop on The Role of Emotion in Human-Computer Interaction, held at the HCI 2005 Conference, Edinburgh, September 2005. http://www.emotion-in-hci.net/workshopHCI2005/positionpapers2005.html
[6] Lindgaard, G., Fernandes G., Dudek C., Brown J. (2006). Attention web designers: You have 50 milliseconds to make a good first impression! Behaviour & Information Technology, Volume 25, Number 2.
[7] Heylen D. , Vissers M. , op den Akker R., Nijholt A. (2004). Affective Feedback in a Tutoring System for Procedural Tasks. Proceedings of the Tutorial and Research Workshop on Affective Dialogue Systems, Lecture Notes in Computer Science 3068, Springer-Verlag Berlin, Heidelberg, New York (2004).
[8] Peter, C., Blyth, G. (2005). The Role of Emotion in Human-Computer Interaction, Workshop at HCI 2005, Edinburgh, UK, 6 September 2005, Volume 2, Pages 295-298, http://www.emotion-in-hci.net.

Computer-Assisted Recording, Pre-Processing, and Analysis of User Interaction Data

Willem-Paul Brinkman
School of Information Systems, Computing and Mathematics, Brunel University, Uxbridge, Middlesex UB8 3PH, United Kingdom
willem.brinkman@brunel.ac.uk

Philip Gray and Karen Renaud
Department of Computing Science, University of Glasgow, G12 8QQ Scotland
{pdg, karen}@dcs.gla.ac.uk

Although computer-assisted recording, pre-processing and analysis of user interaction behaviour has received continuing research attention over the years, its full potential as a data source to inform the design process seems still unrealised. With technologies as broadband internet and distributed applications, it is possible to continuously and unobtrusively collect interaction data. However, what is the best way to record and collect interaction data? What kind of computer tools, i.e. algorithms, can we use to filter and separate relevant data from noise? And also, which types of analyses and measures give us a design-relevant insight into the interaction, the users, their interaction problems, their needs, personality, and experience? Traditionally, psychologists, usability experts, ergonomists, etc. have been among the main consumers and users of this type of data and supporting tools. However, making the data and tools easy accessible to designers and software engineers might even more directly impact the quality of the application. This workshop plans bringing together individuals with an interest in computer-assisted analysis of interaction data with the aim of improving the interaction design. The workshop will be a platform to discuss new ideas and to share experiences, but also to identify new research challenges and potential solutions. Please note that web usage mining in relation to product sales strategies is not within the scope of this workshop.

Sequential data analysis, logfile analysis, computer-assisted usage analysis, user events

1. INTRODUCTION

Using the computer to record and analyse user-system interaction is not new. However, with technologies such as broadband in combination with distributed applications, developers now have access to data sets that are: larger in size as they can collect data over a longer time period; more diverse as they can include users from different geological locations; more up to date as the data is almost instantaneously accessible; and more detailed because of increased availability of storage capacity both locally and centrally. The data itself can be used for a variety of things, such as understanding the usability of the system, the systems functionality usage, or the users' needs, personality, and experience. In short, interaction data is and will become an important information resource for developers and researchers. Where other methods only provide subjective results, such as obtained by questionnaires, interviews, and diary studies; or a snap shot, such lab observation or site studies; online tracking of user interaction provides objective data in an unobtrusive manner, which is ideal for longitudinal studies. Despite these clear benefits, it also has a number of unresolved issues. They relate to the way in which interaction data is recorded, collected, pre-processed and finally analysed. This workshop aims to bring together people from the HCI community to discuss these issues and explore potential solutions.

1.1 Recording and collecting usage data

The kind of data and the way it is recorded and collected is not standardised and currently seems very much a matter of individual taste. Some researchers simply record keypresses or cursor movement with a time stamp, whereas others record higher-level or application relevant events 1. Others have focused on physiological data or other type of behavioural data such as eye gaze. What is recorded seems to relate to the specific implementation of the recording mechanism. For example, if this is implemented in a layer close to the operating system, the recording can include very low-level events 2 such as pressing a single key, or mouse button, or user events that can be understood in the context of the operating system such as creating or killing a process, swapping between applications, opening, resizing, and closing of windows. On the other hand if it is implemented in the application, user events can be recorded in the context of the application, for example a simple press on the enter key becomes a confirmation that a task sequence is completed. A key success factor in the uptake of data collection methods, however, is the ease with which a developer can install this recording mechanism. Several ways seem possible to support developers, such as: providing low-level recording mechanisms tools that are application independent e.g. tracking web access; recording mechanisms that are embedded in the development environment which can simply be activated or deactivated in order to debug a program; or tools for inserting recording routines in specific section of a program 3. Often tools are still in an experimental phase without any large-scale use, and it is unclear whether any standardisation is possible or desirable. Besides the recording, another important issue is the collection of the data. Is this first stored locally and transferred centrally in the computers' downtime? Is there local buffering or is the recording directly transported to a central server? The opposite is also possible, with only local recording and no automated collection facilities, and therefore completely depending on the user or the developer to activate the mechanism and submit the data.

1.2 Pre-processing usage data

Once data has been collected, it often needs to be pre-processed. This could mean 4: 1) selection, separating relevant user events from the 'noise' events; 2) abstraction, relating low-level events with higher-level concepts; and 3) re-coding, generating new event streams based on the results of a selection and abstraction process. Pre-processing seems especially relevant when people are interested in high-level concepts, but only have access to low-level data. Of course this opens the debate about whether we should make high-level tracking easier or whether we should make the tools for pre-processing more powerful.

1.3 Analysing usage data

The final phase is the analysis of the sequential data. Various analysis methods have been proposed 5, such as: Markov analysis, lag sequential analysis, probabilistic finite state machines, Fisher's cycles, maximal repeating patterns, and regular expressions. Whereas these methods are data driven, it is also possible to analyse the data within the context of a model, such as a task model 6, or system-interaction model 7, user models. However, it is unlikely that all potential information about the interaction, the users, their needs, their experiences, their personality etc can already be derived from the data by means of these methods 2. New analysis methods and supporting tools that can help people to analyse sequential data seems therefore required.

2. WORKSHOP FOCUS

The workshop will focus around the three main topics, which are recording and collecting; pre-processing; and analysis. Both theoretical and practical advances, such as tool development, in these areas can be discussed. The workshop will also consider the users of these tools and methods. Traditionally the target group seems to be highly trained experts such as psychologists, usability experts, ergonomists, etc. However, regarding designers and software engineers as the target user group might increase the impact this data can have on the development of applications. The workshop intends attracting both researchers and practitioners that are interested in sequential data analysis. They could have various backgrounds such HCI, psychology, design, or software engineering. The main objective of the workshop is to establish a community of researchers with an interest in this area, allowing a lively exchange of ideas, and a joint exploration of outstanding problems, but also solutions.

3. WORKSHOP FORMAT

This is a full-day workshop. The morning session will consist of short presentations and discussion of participants' position papers. Participants will also have the opportunity to demonstrate computer-assisted tools they might use or have developed. In the afternoon session, participants will break into small groups depending on their main interest and discuss research questions, unsolved problems, potential solutions and new research directions. At the end of the workshop, the small groups will report back, which will form the basis for a plenary discussion. The workshop will also be supported by a website (http://disc.brunel.ac.uk/HCI2006trackingworkshop) hosted by Brunel University. Position papers will be posted here in advance, and participants' slides will also be made available on the website before the workshop. The intention is also to transform the output of the workshop into a call for a special issue of a journal.

REFERENCES

1 Hilbert, D.M. and Redmiles, D.F. (2001) Large-scale collection of usage data to inform design. Proceedings of Interact'01, Tokyo, 9-13 July, pp. 569-576. IOS Press, Amsterdam.
2 Renaud, K. and Gray, P. (2004) Making sense of low-level usage data to understand user activities. Proceedings of SAICSIT '04, pp. 115-124.
3 McLeod, I., Evans, H., Gray, P. and Mancy, R. (2005) Instrumenting bytecode for the production of usage data. In Jacob, R., Limbourg, Q. and Vanderdonckt, J. (eds), Computer-aided design of user interfaces IV. (pp. 185-195). Kluwer Academic Publishers, Dordrecht.

4 Hilbert, D.M. and Redmiles, D.F. (2000) Extracting usability information from user interface events. ACM Computing Surveys, **32**, 384-421.
5 Sanderson, P.M. and Fisher, C. (1997) Exploratory sequential data analysis: qualitative and quantitative handling of continuous observational data. In Salvendy, G. (ed), Handbook of human factors and ergonomics (2ed, pp. 1471-1513). Wiley-Interscience, Chichester.
6 Lecerof, A. and Paternò, F. (1998) Automatic support for usability evaluation. IEEE Transactions on Software Engineering, **24**, 863-888.
7 Brinkman, W.-P., Haakman, R. and Bouwhuis, D.G. (2005) Empirical usability testing in a component-based environment: improving test efficiency with component-specific usability measures. Proceedings of EHCI-DSVIS'04, Hamburg, 11-13 July, pp. 20-37. (LNCS 3425) Springer.

HCI, the Web and the Older Population

Joy Goodman
Engineering Design Centre,
University of Cambridge,
Cambridge CB2 1PZ, UK
jag76@cam.ac.uk

Anna Dickinson
Division of Applied Computing,
University of Dundee,
Dundee DD1 4HN, UK
adickinson@computing.dundee.ac.uk

Suzette Keith and Gill Whitney
School of Computing Science,
Middlesex University, The
Burroughs, Hendon, London
NW4 4 BT, UK
{s.keith, g.whitney}@mdx.ac.uk

Workshop website: http://www-edc.eng.cam.ac.uk/~jag76/hci_workshop06

This workshop on HCI and the older population will provide a forum for the presentation of current work and a platform for discussing key challenges in this area. It aims to build up and support the research community in this area, encouraging communication with older people and addressing key issues together. HCI issues for older people are extremely important in light of the rapidly ageing population in developed countries and the potential offered by computerised technology, a potential that can only be fulfilled if the technology can be used effectively by its target user group. The workshop continues a successful series held at HCI in recent years and this year focuses on the creation and use of websites, a particularly important and timely topic as the use of the internet becomes increasingly integrated with daily life.

1. TOPIC

The population of the developed world is rapidly ageing [3], with implications for many areas of policy and practice, including that of the design of computing and interactive systems. Such systems are becoming more and more a part of daily life and present valuable opportunities for supporting people with age-related impairments. However, older and disabled people often find them unsuitable or difficult to use, due to differences in experience, abilities and situations. The resultant HCI challenges have yet to be comprehensively addressed.

This workshop aims to build up and support the community in the area of HCI for older people, providing a forum for the discussion of important issues, promoting discussion of this topic on a wider level and drawing together an often scattered research community. It builds on previous popular workshops held at HCI in 2002, 2004 and 2005 (*e.g.,* [1]).

1.1 The Internet

In particular, the workshop will focus on how older people create and use websites. Websites are changing from primarily information resources to tools of daily living, and it is therefore increasingly important to consider older people in their design and use. Existing research on the use of the internet by older people has focused on the practical issues of access and usability (e.g., [5]). Beyond these issues, however, is the essential question of why real older people use the internet. Many are now using it for practical purposes and fun by socialising with others and creating personal websites, taking advantage of the opportunities to find others who share their experiences and interests (see, for example, [3,4]).

The afternoon part of the workshop will therefore use the experience of older silver surfers to enable designers and ICT professionals to understand and fulfil the desires of older people for enjoyable web experiences.

1.2. Discussion Topics

The workshop will explore some of the benefits and problems that older people experience when they interact with the web, and will use this as a springboard for the discussion of topics that are relevant not just to web design but also to the whole topic of HCI and the older population. We aim to address questions such as:

- What problems do older people typically have in interacting with technology? How can we help to overcome these?
- What do older people want from technology? Why should they use applications such as the internet?
- What requirements gathering, design and evaluation methods are suitable for use with older people and how can they be used most profitably?

- What can researchers working in widely different application areas, such as web design and Smart homes, learn from each other? What kind of collaboration is possible and useful?
- How can we ensure that our research results have a real impact on the commercial world and on users' daily lives?

2. WORKSHOP PROCEDURE

2.1 Advance Participation
We will ask those interested in attending the workshop to submit short papers (2 pages) on their work so that these can be circulated in advance of the meeting and published as a set of proceedings. This will enable attendees to gain a picture of each other's interests before the meeting and will also provide a record of the workshop for other interested parties. Interested participants who have not submitted a position paper may be considered in special instances but not in the general case.

2.2 Programme
The programme for the day is likely to follow the following format:

Ice-breaker: We will start with an ice-breaker event to encourage the group to mix and get to know each other on an informal basis.

Keynote 1: A short keynote talk will introduce the topic of HCI and the Older Population, describing key issues in HCI and the Older Population and stimulating discussion. This will be given by Roger Coleman, Professor of Inclusive Design at the Royal College of Art and Co-director of the Helen Hamlyn Research Centre, which focuses on the design implications of social change, particularly ageing.

Poster and demo session: A poster and demonstration session will allow participants an opportunity to present and discuss their work and to hear about other research in the area. All participants who have submitted a paper will be invited to display a poster and we are keen for those who have developed products and prototypes to show them to the group. The session may start with a brief introduction in which poster presenters point out their posters and introduce themselves. Members of the older population will be on hand to provide feedback if requested. Posters will also be displayed throughout the day to allow plenty of time for participants to examine them.

Discussion Groups 1: Small groups will then be formed to discuss some of the topics mentioned above. To facilitate this discussion, each group may be given specific questions or tasks to look at. After a time of discussion, the groups will report back to the workshop as a whole, summarising their key points and conclusions.

Keynote 2: The afternoon session will be opened by a short keynote address from a member of the older population from Hackney Age Concern Silver Surfers, introducing the Silver Surfers group and creating personal websites.

Design Workshop: We will then break into smaller groups, each of which will be asked to review or design a game that will entertain an older person and support intergenerational play. Invited older participants from Hackney Age Concern Silver Surfers will be available to contribute to the design process. The groups will share their resultant designs and lessons learned from their older participants. This session will last for about 2 hours altogether.

Discussion Groups 2: Following on from the design workshop, the groups will discuss how we can encourage researchers and designers to respond to the real desires and pleasures of older people.

Wrap-up: As a group as a whole, we will look back on the workshop and try to extract key points from the discussions, focusing on what work needs to be done in this area and how to follow up the workshop. Before finishing, the organisers will sum up the findings of the day and we hope that the group will go out for a social activity in the evening as a further step towards building a community.

2.3 After the Workshop
The submitted papers will be published in dedicated proceedings in electronic and printed form. Following on from the example of the previous workshops in this series (*e.g.,* [2]), we also plan to publish a journal special issue, with *Universal Access in the Information Society*, with submissions solicited from workshop attendees.

REFERENCES

[1] Goodman, J. and Dickinson, A. (eds.) (2005) HCI and the Older Population, Workshop at HCI 2005, Edinburgh, UK, 5 Sep 2005 (http://www.computing.dundee.ac.uk/projects/kit/British_HCI_2005/Workshop.asp).
[2] Goodman, J. and Lundell, J. (eds.) (2005) HCI and the Older Population. Special issue of Interacting with Computers 17(6), Dec 2005.
[3] Hackney Silver Surfers (2006) http://hackneysilversurfers.org.uk/themes/Senior_Citizens/Silver_Surfers. Updated Apr 2006
[4] ICM Research (2002) Internet turns on men and women in different ways - new survey reveals. http://www.icmresearch.co.uk/reviews/2002/it-internet-old-people.htm. Accessed: 25 Apr 2006.
[5] Milne, S., Dickinson, A., Carmichael, A., Sloan, D., Eisma, P. and Gregor, P. Are guidelines enough? An introduction to designing web sites accessible to older people. IBM Systems Journal 44(3), 2005, pp. 557-571.
[6] U.S. Census Bureau (2005) International Data Base (IDB). http://www.census.gov/ipc/www/idbnew.html, updated Apr 2005.

Graduate Career Development Workshop for Women in HCI Research

Lynne Hall
School of Computing & Technology, University of Sunderland, Sunderland, SR6 0DD, UK
lynne.hall@sunderland.ac.uk

Ursula Martin
Department of Computer Science, Queen Mary University of London, Mile End Road, London E1 4NS, UK
ursula.martin@dcs.qmul.ac.uk

This workshop is intended for graduate women in Human Computer Interaction research and aims to bring together a broad range of female researchers and academics to explore career development issues. The workshop addresses the issue of under-representation of women in UK computing research and particularly in UK computing research leadership. The workshop will involve panels and presentations by female research leaders, aiming to share their experiences and successes with participants at the workshop. A range of topics will be considered including issues related to developing and establishing a research area and team, becoming a research leader and the current climate for research. Issues related to work/life balance for women researchers in computing will also be explored. The workshop will be held in conjunction with Women@CL and a key outcome will be to provide participants with female role models in HCI research and networking contacts for future personal development.

Women in computing, gender, leadership, career development

1. INTRODUCTION

The number of women in academic leadership roles in technology is small, and women are under-represented at all stages in UK computing research. This situation worsens throughout the research pipeline, with the proportion of women declining at each career milestone [2]. Only 1 in 4 computing PhDs, 1 in 8 computing academic staff and 1 in 20 computing professors are females. Recent data presented in the OST Athena Asset Survey of male and female university scientists showed that women, even at senior levels, felt disadvantaged in matters such as social inclusion and access to career development, yet 33% of academic women, as opposed to 22% of men, aspire to leadership positions [4]. HCI has traditionally attracted a higher number of women than most other areas of computer science, with increasing recognition that including more HCI in the curricula of computer science courses will attract and retain more women in computing [3]. However, similar to other areas in computing, women are still under-represented in leadership roles in HCI research.

To redress this balance, developing strategies to support academic life and to enable success is vital [1]. The Office of Science and Technology (OST) Greenfield Report [5] identifies that the greatest obstacles to the progression of women already launched in the high-tech field are lack of the support mechanisms such as mentoring, role models, and access to informal networks that support career progression [7]. This highlights a clear need for female-centric professional development opportunities [6] that foster the establishment of role models and networking.

The aim of this workshop will be on understanding women and leadership, particularly how women can make themselves more effective leaders. Participants will leave with a greater self-awareness, new ideas, and information on steps they could take to advance their leadership performance to ensure their progression. The aims for this workshop are for participants to depart with:

- New thoughts and ideas about leadership, women in leadership, and themselves as research leaders
- Greater understanding of the development of research areas and teams
- Appreciation of the political framework for current research
- Role models and networking contacts for future personal development

This workshop will be held with the national Women@CL network. This network provides local, national and international activities for women engaged in computing research and academic leadership. Women@CL's purpose is to put into place a positive action program to encourage and support women in computing research. In 2005, Women@CL held a successful workshop at the International Joint Conference on Artificial Intelligence (IJCAI). Here, the intention is to continue this tradition, with this workshop focused at women researchers in HCI.

This workshop will provide the opportunity for participants to hear successful women in HCI research speak on topics relevant to their careers. It will also give participants a chance to network with their peers and engage in an exchange of views and ideas which foster growth.

This workshop aims to provide a leadership and career planning workshop for early to mid career female researchers in HCI and will focus on a number of specific themes that have particular resonance for female researchers in HCI:

- the nature of research leadership
- choosing an independent research area and building a team
- the impact and potential of the UK Research Assessment Exercise
- opportunities offered by European and trans-national collaboration
- juggling work-life balance

2. WORKSHOP PROCEDURE

The workshop will be divided into the following sessions:

- **Welcome**
  Ursula Martin, Lynne Hall: Overview of the workshop and introduction, information about women@CL
- **Plenary session:** Eileen Scanlon: Experiences in research leadership: developing a research career
- **Panel session: Becoming a leader in HCI research**
  In this panel, we will discuss the nature of research leadership, choosing an independent research area and building a team, and juggling work-life balance. Panel members include Dianne Murray, Eileen Scanlon
- **Plenary session:** Linda Hole: Snakes and the Career Ladder: a grumpy old woman's view
- **Parallel panel sessions**
  - ***Research Issues at early career stage*** (aimed at researchers completing PhDs, post-docs and new academics: we will discuss topics relevant to early career researchers such as how to develop and manage your research, job opportunities in academia and industry, getting established in a research career, issues about work/life balance. Panel members include Janet C. Read, Kate Hone
  - ***Mid-career research issues*** (aimed at established researchers and academics): we will discuss topics to mid career researchers, such as establishing and sustaining a research team, developing research credentials and credibility, accessing research funding, managing the research process, issues related to work/life balance. Panel Members include Linda Hole, Helen Petrie.
- **Plenary session:** Kate Hone: Gender issues and applying for research funding
- **Panel: The Research Climate**
  In this panel, we will discuss the current research climate for women researchers in HCI, this will include a consideration of the impact of the Research Assessment Exercise, both in the remainder of this RAE period and after 2008; and a discussion of the potential offered by European and trans-national research, along with related work/life balance issues. Panel members include Ann Blandford, Linda Hole, Helen Petrie.
- **Plenary session:** Ann Blandford, Winning friends and influencing people: thinking strategically about the future of HCI

The Call for Participation will be submitted to relevant mailing lists e.g. CHI, British HCI, UKGrad, Women in Computing in the UK and internationally and a micro-site at http://osiris.sunderland.ac.uk/~cs0lha/BHCI-women-workshop.html. This workshop is intended for all women in HCI research. Participants are encouraged to submit a short biography prior to the workshop.

REFERENCES

[1] Adams, B., Jensen, K., Lester, R., Olson, D. M., and Tennant, J., "Coping strategies for women in computer science," Journal of Computing Sciences in Colleges, vol. 21, pp. 76-90, 2006.
[2] Camp, T. and Gurer, D., "The incredible shrinking pipeline," http://www.acm.org/women, 2003.

[3] Computing Research Association, "Broadening Participation in Computing Research and Education: Workshop Report," Computing Research Association October 20–21, 2004 2004.
[4] HESA, "HESA and other data available at http://www.set4women.gov.uk."
[5] OST, "Greenfield report," Office of Science and Technology 2003.
[6] Todd, K., Mardis, A. L., and Wyatt, P., "We've come a long way, baby!: but where women and technology are concerned, have we really?," presented at 33rd annual ACM SIGUCCS conference on User services, Monterey, CA, USA, 2006.
[7] Wellington, S., Marcia Brumit Kropf, M., and Gerkovich, P. R., "What’s holding women back," Harvard Business Review, June, 2003.

Designing the Not-Quite-Yet

Ann Light,[1] Patrick G.T. Healey[1] and Gini Simpson[2]
[1] Interaction, Media and Communication Research Group, Department of Computer Science, Queen Mary University of London, E1 4NS
[2] Head of Media Arts, SPACE, 129 – 131 Mare Street, London E8 3RH
annl@dcs.qmul.ac.uk, ph@dcs.qmul.ac.uk, gini@spacestudios.org.uk

As digital technologies become more complex and further penetrate domestic, educational, social and political areas of life, we ask how people are to make sense of the choices available to them. Ideally, everyone should participate in the design decisions that will impact on their lives, based on a good understanding of the potential of digital networks and the implications of using them. But the benefits and shortcomings of technologies are notoriously difficult to anticipate before implementation, just as their uses are. This workshop is for those interested in empowering the public to contribute to design effectively. We will gather and share ideas and success stories in the company of artists working in this field. The goal is a wider franchise for design… and better outcomes.

Digital networks, people's understanding, interdisciplinary, methods, future challenge

1. INTRODUCTION

How do we broaden the constituency of design? How do we help people engage with transformations engendered by technology? How do we enable the appropriation of an invasive, yet intangible, 'internet of things'? This workshop will explore the potential of innovative methods, such as performance, public art, games, etc, to deliver methods that help people do 'design thinking', widen the design franchise, and base the design of future technologies more closely on society's needs and desires. If the next major design challenge in computing is to incorporate digital networks into the fabric of life [1, 2], we ask: how far can that design process engage with the intended users of these new systems - the public?

1.1 Background

Ambient intelligence, pervasive computing, augmented reality, a network of things, smart buildings and clothes, identity tagging, …the digital future is promised as connectivity 'anytime, anywhere', as seamless flows of information between environments, objects and people [3,4,5,6,7]. More than ever this entails an attempt to inscribe social practices and institutions in technology.

We are seeing rapid growth in experimentation with embedded chips, identity tagging, monitoring and other phenomena in the 'network of things'. Retail chains are experimenting with tracing technologies to make their sourcing more efficient and to map consumer habits in ever greater detail [9]. Civil liberty and consumer advocacy groups have raised concerns over privacy if individual products can be traced even within a person's home [9][10]). The Foresight think tank report "RFID-tagged driverless cars on roads by 2056" [8] informs us that: "The UK's transport infrastructure will be radically changed over the next 50 years by RFID tracking tags, embedded sensors and an artificial intelligence network that will reduce congestion and pollution."

The social challenges arising from these developments have been noted, but Human-Computer Interaction (HCI) specialists have not fully engaged with the question of how to design for them. These challenges are particularly interesting because they require users to think like designers or submit to being run by the software surrounding them. This workshop asks how we can create the competent and empowered 'end-designer' [11].

We would expect methods drawn from the arts, education and science/social science to be relevant and we invite participants to bring them and share them at this event running during HCI2006. So if you work to involve, engage or educate the public on the potential of digital technology – especially in innovative or experiential ways – or you'd like to know more about how to do it, sign up. The emphasis will be on sharing practical work, not formal presentations.

A particular feature of this workshop is the participation of a group of artists commissioned by Space (a London media arts and education charity) under the umbrella of looking at the potential of identity tagging technologies, such as RFID, and communicating this potential to the public. This confluence of human design, technological and political concerns is being subjected to artistic interrogation through these artworks, part of Space' "Tagged" initiative. The artist commissions presented at the workshop will be used as lens through which to examine the following issues:

- What do users need to know? How can this understanding be made available?
- e.g. educational agendas, participatory design, methods of involvement, arts and performance practice
- Digital technologies: empowerment or control?
- eg use of RFID, surveillance and monitoring, incidental interactions, community re-appropriations
- Making your presence felt where it matters –
- eg design for political engagement, sharing stories, controlling your representation.

2. GETTING INVOLVED

This workshop is intended to open discussion on how to engage more people in design in the context of the increasingly dense information space surrounding the ordinary spaces we occupy. We expect methods drawn from art, science and social science to be relevant and hope to welcome newcomers to the conference. If you are interested and/or experienced in the areas of public engagement, participatory design, ambient intelligence, the social impact of digital networks or any other relevant area, please consider submitting a position statement and indicating whether, time permitting, you would like to make a five minute presentation or involve workshop participants in a demonstration of techniques for engaging a wider audience in design. Non-presenting participants will be welcomed as long as they are prepared to participate in activities.

3. WORKSHOP ORGANISERS

Ann Light is a member of the Interaction, Media and Communication Group at Queen Mary University of London and co-runs a campaign called "Transform-Ed" (www.transform-ed.org) on bridging the divide in society's comprehension of the potential of digital networks. She is chair of trustees for a digital media charity (www.fiankoma.org) and edits UsabilityNews (www.usabilitynews.com) on a part-time basis. She was once a drama teacher and still uses this background in her interpretation of 'interaction design'.

Pat Healey leads the Interaction, Media and Communication Group and Augmented Human Interaction Laboratory at QMUL. He is interested in the potential of digital technologies to provide uniquely flexible media which transform human communication. Pat's research applies models of human communication - drawn mainly from psychology and sociology - to understanding these processes.

Gini Simpson is the head of SPACE Media Arts, based in Hackney, East London. SPACE Media Arts undertakes large scale electronic arts projects linking artists and communities and provides open access to new technology in East London. This has included working with award winning artists, street gangs from Bow and patients at a London psychiatric hospital. Previous to this, Gini worked for DDB Advertising and Magic Lantern productions iTV. She has produced art events nationally and internationally, including the production of the first New Media marquee and field at Glastonbury Festival.

REFERENCES

[1] Weiser, M., (1991) The computer of the twenty-first century. Scientific American, 94-141.
[2] Chung, E., Hong, J.I., Lin,J., Prabaker, M.K., Landay, J.A. and Liu, M.A (2004). Development and Evaluation of Emerging Design Patterns for Ubiquitous Computing. In Proceedings of DIS2004.
[3]. Gellersen, H.W. Schmidt, A, and Beigl, M. (2002) Multisensor context-awareness in mobile devices and smart artefacts. Mobile Networks and Applications, 7(5):341-351.
[4] Ishii, H.and Ullmer, B. (1997) Tangible bits: Towards seamless interfaces between people, bits and atoms. In Proceedings of Conference on Human Factors in Computing Systems (CHI '97), 234–241. ACM, March 1997.
[5] Mann, S. (1997) Wearable computing: A first step toward personal imaging. IEEE Computer, 30(2), 1997.
[6] Schmidt, A. and Van Laerhoven, K.(2001) How to build smart appliances. IEEE Personal Communications, 2001:66–71
[7] Perry, M., O'Hara, K., Sellen, A. Harper, R. and Brown, B.A.T. (2001) Dealing with mobility: understanding access anytime, anywhere. ACM Trans. on Human-Computer Interaction, 8 (4), p. 1-25.
[8] McCue, A. (2006) RFID-tagged driverless cars on roads by 2056, Silicon.com, (26.1.2006) http://www.silicon.com/publicsector/0,3800010403,39155962,00.htm
[9] El Amin, A. (2005): The Age of Consent for RFID, article on food& drink europe.com (20.6.2005), http://www.foodanddrinkeurope.com/news/printNewsBis.asp?id=60758
[10] Frost, R. (2006) RFID: Beyond the Barcode, brandchannel.com http://www.brandchannel.com/start1.asp?fa_id=295
[11] Light, A (2002) Users as Designers - Challenge of the Network Age, UsabilityNews (8.5.2002) http://www.usabilitynews.com/news/article433.asp

Tutorials

Forms that Work: Understanding Forms to Improve their Design Full-day Tutorial

Caroline Jarrett
Effortmark Ltd, 11 Bridge Street, Leighton Buzzard, LU7 3AB UK
carolne.jarrett@effortmark.co.uk

Forms are ubiquitous, essential and widely disliked, but they do not have to be a difficult experience for the user. This tutorial concentrates on the human side of forms: how we interact with, and think about forms and how to make them as easy for users as practicable.

You do not need any special knowledge or experience of forms or usability – but if you do, you'll enjoy comparing your ideas about forms with mine. You need to have some interest in forms and preferably to be working with forms in some way.

This is in the 'tutorial' format, but you won't be sitting there to be lectured at. You'll be working on practical exercises that contribute to a forms process that you can take away to use immediately.

1. WHAT YOU WILL LEARN

You will learn:

1. The three-layer model of the form: relationship, conversation and appearance
2. A six-step process of creating forms that work for the users
   - Define the information requirement
   - Understand the relationship
   - Write questions
   - Create a conversation
   - Design an appropriate appearance
   - Testing

   The importance of iteration throughout the process.

Take-aways include:

- a deep understanding of forms, encapsulated in a variety of guidelines that are easy to remember and apply
- thorough appreciation of the three-layer model of forms (relationship, conversation and appearance) and how to use it to improve forms
- a range of ways of testing forms and understanding of the value of testing.

2. WHAT TO EXPECT

The style of the day is very interactive. Each section includes at least one exercise where you will use some of the ideas that have just been presented. Discussion and questions are welcome throughout.

DETAILED OUTLINE

The information requirement

Error rates on paper forms are frequently extremely high. This section of the day looks at finding out what you need to find out, and finding out what data your organization already holds and uses.

Relationships: the users' reactions to forms
People react differently to forms than to other pages on a web site or in a package of documentation. This section discusses the relationship changes from box to box as the user works through the form. We also look at trust and research on response rates in questionnaires as a basis for thinking about response rates / dropout rates on forms.

Conversation part 1: Questions
This topic looks in detail at how users answer questions, or find or construct the answer to a question. We look briefly at the cognitive aspects of reading, then go into more detail on 'locating the answer' and how to choose the appropriate response spaces (also known as 'controls').

Conversation part 2: topics and flow
A great form is like a pleasant conversation: you hardly notice that you've given out information because you're enjoying it. This section looks at creating a flow across topics, designing validations, and sorting out a preamble (also knows as the instructions at the start of a form).

The appearance of forms
If forms look good, they are nicer to fill in. This section is about designing an attractive appearance for the form. Throughout this topic, participants will redesign a typical log-on form to create a paper prototype that they test in the final session of the day.

Testing
The final session of the day looks at layer checks (desk checks that relate to the three-layer model of the form that underlies the construction of the day) and then at usability testing. You will try out appropriate checks on their newly designed prototypes.

We then look at how to adapt usability testing for testing forms, and you will try a short test of the prototype forms.

The day closes with a brief review of the key topics and a final opportunity for questions.

Principles of Interaction Design

Shane Morris
Echo Interaction Design, Australia
shane@echointeraction.com.au

Great interaction design requires more than just a methodology. Through years of practice, master interaction designers build up a body of design 'intuition' that allows them to rapidly design excellent user interfaces.

This tutorial introduces and explores many of the fundamental principles that underlie the practice of interaction design and user interface design. It is by consciously or unconsciously drawing on these principles that practising designers are able to efficiently produce excellent designs, and minimise redesign. Nevertheless, apart from some 'heuristics', many practitioners have not been formally introduced to these principles.

This highly interactive tutorial gives participants a 'kick start' towards master status by exploring the underlying principles of interaction design through exercises, discussion and examples. Practicing interaction designers will develop their working vocabulary of design principles that can be applied in the design and evaluation of all forms of user interfaces in everyday use - including physical devices, graphical user interfaces and internet applications.

User interface design, interaction design, design principles, design thinking

1. TECHNIQUES, PRINCIPLES AND INTUITION

In all design disciplines, student designers are initially introduced to a set of techniques to help them analyse design problems, generate design ideas and evaluate their solutions. These formal methods allow the beginning designer to be productive, before they have developed their own design intuition. Beginning designers follow predictable and measurable steps intended to increase the likelihood of producing a satisfactory design, within the prescribed timeframe. Note, however, that merely practicing these techniques does not guarantee an excellent solution.

At the other end of the spectrum, the difference between a 'master' designer and a merely competent one is the master designer's development of a design 'intuition' which allows him or her to quickly understand a problem and to efficiently generate one or (more likely) more high-quality solutions. It is understood that this 'intuition' is difficult for the master to verbalise, indeed attempts to systematise such 'expert' knowledge are often unsuccessful.

Between formal techniques and design intuition lie design principles. Design principles do not belong to one particular technique or another, but are applied across the design lifecycle - consciously and unconsciously - to facilitate the creation of excellent, timely design. Design principles are less prescriptive than methodologies, in that they are often applied in a less structured way, even unconsciously. Design principles identify 'best practices' at a level of abstraction where they can be applied to a range of design problems. As such, they provide an important stepping-stone to the development of a designer's own design intuition.

Intuition (wisdom)

Principles (e.g. progressive disclosure)

Techniques (e.g. heuristic evaluation)

3 progressively more sophisticated types of design knowledge.

2. THE IMPORTANCE OF HAVING PRINCIPLES

The field of interaction design benefits from having drawn practitioners from a variety of backgrounds. However, this means that many practitioners come to the field without formal training in user interface design. Many practicing interaction designers have never been formally introduced to these principles of design, or are possibly only familiar with a few of them (in the guise of 'heuristics', perhaps). They apply personal experience, intuition, imitation and extensive evaluation to produce and refine the user interfaces they are responsible for. Along the way, they build up a body of design experience.

This tutorial aims to introduce interaction designers to some key principles of interaction design (and design generally). Through discussion, examples and exercises participants will build their familiarity with and ability to recognise and apply these principles in their day-to-day work. In doing so, participants will build their skills in producing designs which are based on sound principles, are defensible and are of higher quality – all in less time.

2.1 Lingua Franca

As well as assisting in the creation and evaluation of user interface designs, familiarity with these design principles gives participants a practical vocabulary to aid communication with colleagues and other interaction designers in the workplace.

3. PRINCIPLES COVERED IN THE TUTORIAL

In this tutorial we explore 26 key principles of interaction design. They include universal principles like *functional layering* and *visual hierarchy*, fundamental interaction guidelines like *Fitt's Law* and cognitive concepts like *gestalt grouping theories*.

This highly interactive tutorial introduces each principle in turn, along with positive and negative usage examples. Short individual and group revision exercises are used throughout the day, and the tutorial ends in a rapid design exercise which aims to explore the pro's and con's of applying the principles in day to day practice.

4. APPRENTICE TO MASTER

Most importantly, the principles are explored in a practical context. While a working knowledge of the design principles covered in this tutorial helps practitioners keep these issues 'top of mind' when designing and evaluating, and assist in collaboration with other contributors, it is only through extensive practical application that designers become fluent with these principles – to the point of them becoming 'automatic' (or intuition).

All design disciplines acknowledge that formal education can only get a prospective designer half-way to competence, the rest is achieved by building a body of less formal intuitive knowledge. This is achieved not through formal education, but through extensive practice. This is why design schools for all design disciplines incorporate extensive 'studio' work in their education programs. It is through exposure to a range of design problems, through critique and criticism in a safe environment and through trial and error that budding designers hone their skills. It is interesting that the highest profile interaction design schools have embraced this studio culture, but many have still to fully acknowledge this necessary component of the education of a professional designer.

Acknowledging that formal learning can only develop a designer so far, this tutorial includes a large component of practical work. While a day's tutorial cannot hope to reproduce years of studio-based experience, we hope, at least for a day, to create an atmosphere where practitioners can safely explore their design options, away from the glare of clients and deadlines. What is more, discussion arising from the practical exercises provides a much more compelling and directed environment in which to explore the principles in question.

5. WHAT THIS TUTORIAL IS *NOT* ABOUT

This tutorial is *not* about *design process*, but rather the underlying knowledge that designers apply when generating design solutions, regardless of the process.

Neither is this tutorial about *visual design*, although some principles of visual communication are covered.

6. LEARNING OUTCOMES

The main issues to be covered by this tutorial are:

- What are the main interaction design principles underlying the design of effective user interfaces?
- How do the principles support or interfere with each other?
- How can the principles be applied in the practical design and evaluation of user interfaces?

Participants will gain knowledge and skills in: ("Learning outcomes")

- Familiarity with the principles of interaction design, their applicability and the relationships between them.
- Skill in applying the principles of interaction design to the evaluation of their own or others' design work.
- Practical techniques to apply principles of interaction design in generating design solutions, and in discriminating between design options.

Old Cards, New Tricks: Applied Techniques in Card Sorting

William Hudson
Syntagm Ltd, 10 Oxford Road, Abingdon, OX14 2DS, UK
william.hudson@syntagm.co.uk

Card sorting is an extremely useful technique in the design of interactive systems. However, it is under-used in practice – often through a lack of understanding and the complexities of cluster analysis. This half-day, hands-on tutorial uses concrete examples taken from live web sites to guide participants through the analysis, design and execution of card sorting activities, particularly as they apply to web navigation. Specific topics presented include open and closed sorting, rapid data collection using bar codes, cluster analysis and extensions to traditional analyses using quality of fit metrics and measures of deviation.

Card sorting, cluster analysis, dendograms, dendrograms

1. INTRODUCTION

(This brief paper provides a flavour of the content of the tutorial. It is based on a longer article written for ACM *interactions* [1]).

Card sorting is a knowledge elicitation technique often used by information architects, interaction designers and usability professionals to establish or assess the navigation hierarchy of a web site. The items are typically menu entries or hyperlinks while the groups are categories or headings. The process involves asking participants to sort items into meaningful groups. In open card sorts the number and names of groups are decided by each participant while in the closed card sorts these factors are fixed by the researcher in advance.

2. CARD SORTING ANALYSIS

Analysis of card sorting results range from simple counting of the number of times items were grouped together to the rather intimidating monothetic agglomerative cluster analysis (known simply as cluster analysis in most cases). Unfortunately, no single technique provides everything a researcher needs to know, especially if convincing evidence is needed to persuade colleagues or customers of the effectiveness of a proposed design.

2.1 Convincing Evidence

The evidence we need falls into three categories:

- Participants. Are these the right participants for our site? Are they all thinking about the items and their groupings in a similar way? Do they have a clear understanding of the card sorting task itself?
- Items. Are the item names well-understood by participants? Are there alternatives that should be considered – perhaps terms users are more familiar with?
- Groups. For closed card sorts, have we chosen the right number of groups and names for each? For open sorts, are participants largely in agreement about the number of groups needed? How well do participants feel the items fit into their groups?

Happily, the answer to this last question – how well participants feel the items fit into their groups – can also help us with many of the other issues listed. Coupled with a few data collection guidelines and alternative presentations of results, we can collect fairly comprehensive evidence about what is and what is not going to work in our navigation hierarchies.

2.2 Quality of Fit by Item

So let's examine this last question in some more detail: How well do participants feel the items fit into their groups? It is possible to argue that this question is redundant; that the items must fit into their groups relatively well in any given set of results, because that is how the participant decided to group them. However, practical experience says otherwise. In many cases, participants place items into groups that are "good enough" but not necessarily ideal (satisficing versus maximizing behaviour). By asking participants to provide a "quality of fit" measure the cluster analysis can be extended to include the strength of relationship between items. Furthermore, analyzing quality of fit by participant or by item can give us useful insights into the some of the thought processes behind the sorting activity. For example, *figure 1* shows quality of fit averaged by item[1] for a closed card sort based on wines. The highest quality of fit in this example is Claret while the lowest is Beaujolais. They are both red wines so it seems a little curious that the graph shows such disparity. The answer lies in the group names that were assigned in advance: full-bodied reds, dry whites and sparkling. Participants were happy to put both Claret and Beaujolais into "full-bodied reds" even though they apparently knew that Beaujolais was not full-bodied and despite instructions that they were to omit items they felt did not fit into any group.

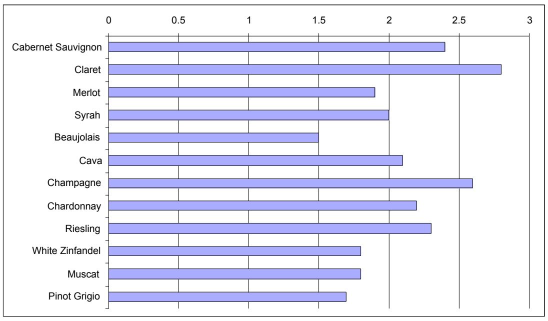


FIGURE 1: Mean quality of fit by item

In this case the quality of fit metric has allowed participants to express their dissatisfaction with Beaujolais (in the full-bodied reds group) while acknowledging their reluctance to discard the item entirely.

2.3 Alternatives to the Dendogram

Card sorting results are traditionally shown using a tree-like figure known as a dendogram. Unfortunately, in many applications, dendograms hide more information than they show. A simple alternative that is explored in the tutorial is a simple surface map of the underlying proximity matrix used for cluster analysis (each cell in the matrix is the frequency with which card pairs appeared in the same groups weighted by average quality of fit for the pair). An example, produced using Microsoft Excel, is shown in *figure 2*.

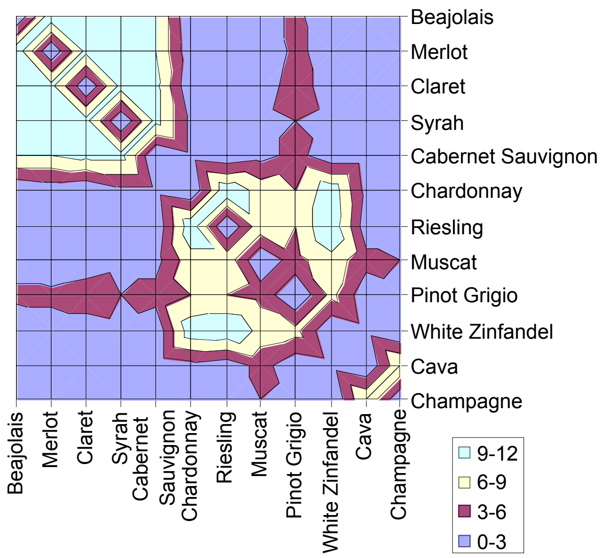


FIGURE 2: Surface map of the weighted proximity matrix

The map shows three distinct groups with anomalies centred on Pinto Grigio. Further investigation brought to light a terminological confusion: participants were not generally familiar with Pinot Grigio (white) and were associating it with Pinto Noir (red). Consequently, participants grouped Pinto Grigio with red wines almost as frequently as they did white. While a hint of this problem would have been discernable in a traditional dendogram, it would not been as dramatically obvious.

REFERENCES

[1] Hudson, W. (2005) Playing your cards right: getting the most from card sorting for navigation design. *interactions*, **12(5)**, 56–58.

[1] Quality of fit is indicated by participants on the following scale: fair (1), good (2) or perfect (3). If an item is omitted by a participant, its quality of fit is 0.

Interviewing Skills for Usability Testing
A Half-Day Tutorial

Caroline Jarrett
Effortmark Ltd, 11 Bridge Street, Leighton Buzzard, LU7 3AB UK
caroline.jarrett@effortmark.co.uk

If you've learned usability testing by practicing on fellow-students: did you miss anything? What happens when you try those techniques on people who are different from them?
If you've learned usability testing from books: did you feel the need of a chance to practice? This tutorial is for you if you have run between 1 and 5 usability tests and want to practice your interviewing techniques. We'll concentrate on the actual interviewing part: asking questions, persuading your participants to think aloud, offering help when appropriate.

If you're an experienced usability or HCI professional then you may find that you already know much of the material presented here – but you may also find it helpful to compare your ideas with mine.

WHAT WILL YOU LEARN?

You will:

- Understand and use six basic types of questions in the context of usability tests
- Learn five tips for better listening in usability tests
- Practice and share ideas for coping with difficult participants
- Understand and select the appropriate choices for offering help
- Know how to construct the entry and exit parts of the test

HOW WILL YOU LEARN?

There is a brief PowerPoint presentation that introduces the key ideas for each section and provides references to support the ideas we cover.
Then we launch into practical exercises to try out the ideas. You will take turns as interviewer and participant, working on a selection of paper prototype web sites.

DETAILED OUTLINE

Introduction

Warm up section that introduces the topics to be covered.

The first exercise asks you to interview your neighbour so as to introduce them to the other people in your group. In the analysis of the exercise, we reflect on questions, answers, listening and body language.

Basic questioning and listening

'Basic questioning' draws on Wallace V. Schmidt and Roger N. Conway (1999) *Results-Oriented Interviewing: Principles, Practices and Procedures* Allyn and Bacon, Boston to present five types of question and seven types of probe. We construct a small 'interview' that uses all five types of question, then you practice the questions and probes in an interview.

In 'listening', you are asked to rate yourself as a listener. The key concept is that listening is contextual. This section closes with a discussion of the directive and facilitative styles of interviewing, showing how both styles are important in a usability test.

Participants and their behaviours

On the whole, participants want to be helpful in usability tests. But what do you do when you get one that doesn't want to help, or who wants only to help and won't say anything critical?

In this section, we describe some ways of coping with participants' conscious and unconscious difficult behaviours. Then you practice your coping strategies in a short interview.

This section includes tips for working with people with disabilities.

Think-aloud and offering help
For most usability tests, we want our participants to think aloud as they work. What happens if they don't? And if you offer encouragement by asking questions, will your test spill over into offering too much help?

This section describes four techniques for launching think-aloud, and discusses a selection of ways of refusing help while maintaining rapport with the participant.

In the exercise, you'll practice a selection of ways for answering the 'blame' questions that such as "you must think I'm stupid" often accompany requests for help.

Starting up and closing down
Launching think-aloud is a key part of opening the usability test, but not the only one. In the final section of the tutorial, we look at the interview setting: its psychological/social setting, time dimensions and physical setting, and also the legal and ethical constraints.

How to Combine Requirements and Interaction Design Through Usage Scenarios

Hermann Kaindl
Gußhausstr. 27-29, A-1040 Vienna, Austria
kaindl@ict.tuwien.ac.at

When the requirements and the interaction design of a system are separated, they will most likely not fit together, and the resulting system will be less than optimal. Even if all the real needs are covered in the requirements and also implemented, errors may be induced by human-computer interaction through a bad interaction design and its resulting user interface. Such a system may even not be used at all. Alternatively, a great user interface of a system with features that are not required will not be very useful as well.

Therefore, we argue for combined requirements engineering and interaction design, primarily based on usage scenarios. However, scenario-based approaches vary especially with regard to their use, e.g., employing abstract use cases or integrating scenarios with functions and goals in a systematic design process. So, the key issue to be addressed is how to combine different approaches, e.g., in scenario-based development, so that the interaction design as well as the development of the user interface and of the software internally result in an overall useful and useable system. In particular, scenarios are very helpful for purposes of usability as well.

Interaction design, usage scenarios, requirements engineering, user interfaces, usability

1. PURPOSE

This tutorial is targeted towards people who are supposed to work on the interaction design or the requirements in systems development, e.g., interaction designers, user interface developers, Web designers, requirements engineers, or project managers. Whatever the roles of the tutorial participants actually are in their daily work, they should get a better understanding of “other” viewpoints and tasks and, in particular, a common approach. The overall purpose of this proposed tutorial is to teach how requirements engineering and interaction design relate and how they can be usefully combined. This can be important for creating better interactive systems in the future.

2. KEY LEARNING OUTCOMES

In this tutorial, participants learn about combined (concurrent and intertwined) requirements engineering and interaction design. In particular, participants will understand how scenarios and use cases can be utilized both for requirements engineering and interaction design, though with different emphasis on the level of detail. They will also understand the additional need to specify the functional requirements for the system to be built, even in the context of object-oriented (OO) development. Overall, they will gain a better understanding of early systems design.

3. TOPICS COVERED

3.1 Background

- Requirements
- Scenarios / Use Cases
- Interaction design
- Object-oriented modelling features
- Iterative and incremental development

3.2 Functions / tasks, goals and scenarios / use cases

- Relation between scenarios and functions / tasks
- Relation between goals and usage scenarios
- Composition of these relations

3.3 Requirements and object-oriented models

- Metamodel in UML
- Requirements and objects

3.4 A systematic design process

- Navigation in the metamodel graph
- Partial sequences of steps
- Improvements through this process

3.5 Scenarios / use cases for interaction design

- Interaction tasks
- Abstract use cases
- From abstract use cases to concrete user interfaces

3. CV OF THE PRESENTER

Hermann Kaindl joined the Institute of Computer Technology at the Vienna University of Technology in Vienna, Austria, in early 2003 as a full professor. Prior to moving to academia, he was a senior consultant with the division of program and systems engineering at Siemens AG Austria. There he has gained more than 24 years of industrial experience in software development and human-computer interaction. He has published four books and more than ninety papers in refereed journals, books and conference proceedings. He is a senior member of the IEEE and a member of the ACM and INCOSE, and he is on the executive board of the Austrian Society for Artificial Intelligence.

REFERENCES

[1] Constantine, L. and Lockwood, L.A.D. (1999) Software for Use. ACM Press, New York, NY.
[2] Kaindl, H. (1997) A Practical Approach to Combining Requirements Definition and Object-Oriented Analysis. *Annals of Software Engineering*, **3**, 319–343.
[3] Kaindl, H. (2000) A Design Process Based on a Model Combining Scenarios with Goals and Functions. *IEEE Transactions on Systems, Man, and Cybernetics (SMC) Part A*, **30**, 537–551.
[4] Kaindl, H. (2001) Adoption of Requirements Engineering: Conditions for Success. *Proceedings of the Fifth IEEE International Symposium on Requirements Engineering (RE'01)*, invited State-of the-Practice Talk, Toronto, Canada, August, pp. 156–163. IEEE.
[5] Kaindl, H. (2005) Is Object-oriented Requirements Engineering of Interest? *Requirements Engineering*, **10**, 81–84.
[6] Kaindl, H. (2005) A Scenario-Based Approach for Requirements Engineering: Experience in a Telecommunication Software Development Project. *Systems Engineering*, **8**, 197–210.
[7] Kaindl, H. and Jezek, R. (2002) From Usage Scenarios to User Interface Elements in a Few Steps. *Proceedings of the Fourth International Conference on Computer-Aided Design of User Interfaces (CADUI'2002)*, Valenciennes, France, May, pp. 91–102. Kluwer Academic Publishers, Dordrecht, The Netherlands.
[8] Kaindl, H., Kramer, S. and Kacsich, R. (1998) A Case Study of Decomposing Functional Requirements. *Proceedings of the Third International Conference on Requirements Engineering (ICRE '98)*, Colorado Springs, CO, April, pp. 156–163. IEEE.
[9] Kaindl, H., Kramer, S. and Hailing, M. (2001) An Interactive Guide Through a Defined Modelling Process. in *People and Computers XV, Joint Proceedings of HCI 2001 and IHM 2001*, Lille, France, September, pp. 107–124. Springer-Verlag, London, England.

Managing Iterative Projects More Effectively: Theories, Methods and Heuristics for HCI Practitioners

Emeritus Professor John Long
UCL Interaction Centre, University College London
31-32 Alfred Place, London, WC1E 7DP
J.long@ucl.ac.uk

Steve Cummaford
IMG Media
McCormack House, 1 Burlington Lane, London, W4 2TH
SCummaford@imgworld.com

Most HCI specialists are involved, in one way or another, with iterative project management (IPM), as opposed to HCI, on a day-to-day basis. However, few specialists have any systematic training or exposure to IPM. Further, market pressures highlight the importance of HCI iterative and adaptive planning and development to meet changing conditions, associated with technology and customer change. This tutorial is intended to fill these gaps. IPM is characterised in terms of its theory, its methods and heuristics to support its practice. Exercises and mini-practicals support the integration of HCI into the heuristics, methods and theory of IPM. In the light of our recent experience, participants' IPM effectiveness, as either managers or as team members, is expected to increase as a result.

Iterative project management, user-centred design, adaptive planning, resources, risks, project teams

NEED

Most HCI system development takes place within the context of the project, whether cash point, web site or complex systems. Such projects are conducted by project managers and carried out by project teams. Most projects are iterative. As a result, HCI specialists are involved with iterative project management (IPM) on a day-to-day basis, including the iteration of the HCI/User-Centred Design (UCD) contributions, for which they are responsible. However, very few such specialists have any systematic training or exposure to IPM. Further, the increasing pressure of the market place, constraining budgets, while shortening deadlines, ensures there is an ever-greater need to manage projects iteratively and to plan adaptively, consistent with both IPM and HCI/UCD constraints. Otherwise, the potential for IPM error will increase too. Changing plans, including those of UCD, within a fixed budget, has become a routine IPM requirement; but one, which is all-too-often carried out ineffectively. This tutorial supports both managers and team members in meeting this requirement, as well as filling the training/educational gap.

LEARNING OBJECTIVES

The learning objects for the participants are threefold:

- To understand the scope; theory; and practice of IPM, as they relate to HCI.
- To become acquainted with, and selectively to practise, UCD methods, in relation to IPM.
- To become familiar with IPM guidance, in the form of heuristics, which they can integrate into IPM and HCI methods.

Following tutorial attendance. Participants should be able:

- To follow up further their interest in IPM via the literature, the web and other sources.
- To make IPM issues explicit in their own UCD activities and to suggest approaches to them.
- To contribute to IPM best practice, within their own HCI projects.
- To become a more effective HCI member of the IPM team.

PARTICIPANTS

The tutorial is of interest to anyone involved in, or indeed concerned by, HCI/UCD, including: human factors and software engineers; project managers; usability specialists; user experience architects; designers;

graphic artists; web services teams; project sponsors; marketers; financers etc. The pre-requisites for the tutorial are either that participants have some experience of IPM projects, either as managers or as HCI/UCD team members, or that they have an interest in HCI and IPM, directly or indirectly, that they would like to develop further.

TOPICS

The tutorial is divided into three parts: IPM Theory; IPM Method; and IPM Heuristics.

1. IPM Theory

The question of what is IPM, as it relates to HCI/UCD, is answered by proposals, concerning: its definition (What is iterative? What is IPM?); its concerns (resources; effort; people/teams; and risks); its discipline: (research; tools; case-studies) and practice (iterative; incremental; evolutionary; and adaptive); and the requirement for IPM, as it relates to HCI/UCD (common managers' problems and common team members' problems). IPM knowledge is reviewed in terms of: Waterfall problems; research; standards; experts; and business cases. IPM practice is characterised by: its types (risk-driven; client-driven;); its methods (planning; development; and Agile); and its development cycle (pre-production; production; maintenance; and evaluation), as concerns HCI/UCD practice.

2. IPM Method

IPM planning and development methods are reviewed. A web site development method is selected for particular address and is used to illustrate the nature of IPM and HCI/UCD practice. However, its general aspects are emphasised. The major phases and stages of the method are described and illustrated. These comprise: pre-production – project clarification; solution definitions; and project specification; production – content; design; construction; testing, launch and hand-over; maintenance; and evaluation. HCI and UCD practice and methods are situated with respect to the IPM method.

3. Heuristics

Heuristics have been culled from a wide range of different sources, both published and experienced. The heuristics are organised around the following topics: project management: planning; iterating; environment; requirements ; and testing. All topics are illustrated and exemplified. Participants practise mapping the heuristics (3) to the methods (2) in the light of theory (1), as they concern HCI/UCD practice.

FORMAT

The important characteristics of the tutorial are as follows;

The tutorial is conducted in a seminar and interactive style. The aim is to elicit the IPM experiences, both good and bad, of the participants; but, in particular, difficulties they have encountered within HCI/UCD projects, either as managers or as team members. Exercises and mini-practicals occur throughout the tutorial and are intended to pull-through the IPM theory, methods and heuristics into HCI/UCD practice. Where possible, the exercises are oriented towards the IPM/HCI difficulties, identified by the participants. In this way, they are able to learn from their own experiences and from those of others.

As described earlier, participants are involved as follows: seminar delivery of materials; interactive discussion, particularly of participants' IPM/HCI experiences; question asking and answering both from and to the tutors; conduct of IPM/HCI exercises and mini-practicals, both as individuals and in small groups; reporting back to the tutorial group as a whole.

CONCLUSION

The tutorial is new as a conference tutorial (presented for the first time at HCI 05). However, it is based on an IPM module, embedded in an HCI/UCD course, conducted with success for a major UK government department's web services team. In our experience, the tutorial meets the needs set out earlier. As the tutorial progresses, participants come to see their own IPM/HCI experiences in a more insightful light. Further, they get to practise useful skills of diagnosing IPM/HCI problems and contributing, either as managers or as team members, to IPM/HCI solutions. In this way, the IPM/HCI effectiveness of the participants is enhanced.

Index of authors

A

B

H

I

J

O

P

Q

SPONSORS

ISSN: ISSN 1470-5559; RR-06-07